

*The US Army's Center for Strategy and Force Evaluation*

MEMORANDUM REPORT  
CAA-MR-97-43

**STOCHEM INVESTIGATION OF COSAGE  
SAMPLING (SICS)**

**AUGUST 1997**



**PREPARED BY  
TACTICAL ANALYSIS DIVISION**

**US ARMY CONCEPTS ANALYSIS AGENCY  
8120 WOODMONT AVENUE  
BETHESDA, MARYLAND 20814-2797**

**20081029142**



## **DISCLAIMER**

**The findings of this report are not to be construed as an official Department of the Army position, policy, or decision unless so designated by other official documentation. Comments or suggestions should be addressed to:**

**Director  
US Army Concepts Analysis Agency  
ATTN: CSCA-TA  
8120 Woodmont Avenue  
Bethesda, MD 20814-2797**

CAA Memorandum Reports are used for the convenience of the sponsors or the analysts to record substantive work done in quick reaction studies and major interactive technical support activities; to make available preliminary and tentative results of analyses or of working group and panel activities; or to make a record of conferences, meetings, or briefings, or of data developed in the course of an investigation. Memorandum Reports are reviewed to assure that they meet high standards of thoroughness, objectivity, and sound analytical methodology.

<b>REPORT DOCUMENTATION PAGE</b>			<i>Form Approved</i> <i>OPM NO. 0704-0188</i>	
Public reporting burden for this collection information is estimated to 1 hour per response, including the time for reviewing instructions, searching existing data sources gathering and maintaining the data needed, and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information. Including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of information and Regulatory Affairs, Office of Management and Budget, Washington, DC 20503.				
1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE August 1997		3. REPORT TYPE AND DATES COVERED Final, February 1997 - August 1997
4. TITLE AND SUBTITLE STOCES Investigation of COSAGE Sampling			5. FUNDING NUMBER  N/A	
6. AUTHOR(S) Mr. Walter J. Bauman				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army Concepts Analysis Agency 8120 Woodmont Avenue Bethesda, MD 20814-2797			8. PERFORMING ORGANIZATION REPORT NUMBER CAA-MR-97-43	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Director, CAA 8120 Woodmont Avenue Bethesda, MD 20814-2797			10. SPONSORING/ MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE A	
13. ABSTRACT (Maximum 200 words)  The Stochastic Concepts Evaluation Model (STOCES) is a fully automated simulation of theater combat applied by the US Army Concepts Analysis Agency. The STOCES was designed as a stochastic version of the deterministic Concepts Evaluation Model IX (CEM IX). The STOCES uses weapons effectiveness data, denoted as combat samples, which are preprocessed by the Combat Sample Generator (COSAGE) simulation. In the SICS analysis, STOCES case variations are constructed, using a single scenario, with an objective of reducing divergence from deterministic CEM, and of increasing variability of STOCES results over replications. The STOCES case variations are constructed by varying the ways that COSAGE combat samples are used in STOCES, and by varying the types of combat processes represented stochastically in STOCES. For each case variant, assessment is made of STOCES divergence from deterministic CEM, and of variability of STOCES results over replications. Closeness of STOCES results to historical results is also assessed for selected cases. The principal finding is that the preferred method for applying STOCES is to deactivate the stochastic FEBA move rate process and to apply a single randomly selected (in each replication) set of COSAGE samples to all simulated battles in a replication.				
14. SUBJECT TERMS  combat simulation, simulation, CEM, STOCES, COSAGE, stochastic, stochastic simulation, combat samples, stochastic sampling, historical data			15. NUMBER OF PAGES  83	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT  UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE  UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT  UNCLASSIFIED	20. LIMITATION OF ABSTRACT  SAR	



## CONTENTS

		Page
	<b>SUMMARY .....</b>	<b>v</b>
	<b>SICS Briefing.....</b>	<b>1</b>
<b>APPENDIX</b>		
A	RAA Contributors.....	A-1
B	Request for Analytic Support.....	B-1
C	References .....	C-1
D	History vs STOCER Mean Combat MOEs in Each 4-day Period .....	D-1
E	History vs STOCER Mean FEBA Position MOEs in Each 4-day Period .....	E-1
<b>GLOSSARY .....</b>		
		Glossary-1

## FIGURES

### FIGURE

D-1	Average US/UK Ammunition Consumed in Each Period.....	D-1
D-2	Average Cumulative US/UK Ammunition Consumed.....	D-2
D-3	Average German Ammunition Consumed in Each Period .....	D-2
D-4	Average Cumulative German Ammunition Consumed .....	D-3
D-5	Average Cumulative US/UK Permanent Tank Losses .....	D-3
D-6	Average Cumulative German Permanent Tank Losses .....	D-4
D-7	Average US/UK Permanent Personnel Casualties in Each Period .....	D-4
D-8	Cumulative US/UK Permanent Personnel Casualties in Each Period .....	D-5
E-1	FEBA Positions on D+4.....	E-1
E-2	FEBA Positions on D+8 .....	E-2
E-3	FEBA Positions on D+12 .....	E-2
E-4	FEBA Positions on D+16 .....	E-3
E-5	FEBA Positions on D+20 .....	E-3
E-6	FEBA Positions on D+24 .....	E-4
E-7	FEBA Positions on D+28 .....	E-4
E-8	FEBA Positions on D+32 .....	E-5



## STOCCEM INVESTIGATION OF COMBAT SAMPLING (SICS)

### SUMMARY

#### **THE REASON FOR PERFORMING THE RESEARCH ANALYSIS ACTIVITY (RAA)**

was that a preferred method needs to be determined for using Combat Sample Generator (COSAGE) samples in Stochastic CEM (STOCCEM) simulation applications at the US Army Concepts Analysis Agency (CAA), including the Kursk Operation Simulation and Exercise (KOSAVE) Study. Previous STOCCEM applications at CAA have frequently exhibited undesirable behavior.

(1) Mean STOCCEM results have frequently and substantially diverged from corresponding results from the deterministic Concepts Evaluation Model (CEM) at CAA.

(2) The range of STOCCEM variation over replications has frequently been excessively small.

**THE RAA SPONSOR** is the Director, US Army Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20814-2797.

**THE SCOPE OF THE RAA** was to apply STOCCEM for the scenario from the Simulation Enhancements from Ardennes Campaign (SEACA) RAA, while varying the use of COSAGE samples and stochastic processes in STOCCEM, so as to reduce divergence from deterministic CEM and to increase variability of STOCCEM results over replications. For each case variant, assessment was made of STOCCEM divergence from deterministic CEM and of variability of STOCCEM results over replications. Closeness of STOCCEM results to historical results was also assessed for selected cases.

**THE MAIN ASSUMPTION** is that 30 replications of STOCCEM applied with the SEACA scenario are sufficient to produce useful results.

**THE BASIC APPROACH** was to devise STOCCEM application cases which varied the types of active stochastic processes and the method for using COSAGE samples and to subsequently assess results from these cases for divergence, relative to deterministic CEM, and variability over replications.

**THE PRINCIPAL FINDING** was that the preferred method for applying STOCCEM is to deactivate the stochastic forward edge of the battle area (FEBA) move rate process and to apply a single randomly selected (in each replication) set of COSAGE samples to all simulated battles in a replication of STOCCEM.

**THE RAA EFFORT** was directed by Walter J. Bauman, Tactical Analysis Division, US Army Concepts Analysis Agency

**COMMENTS AND QUESTIONS** may be sent to the Director, US Army Concepts Analysis Agency, ATTN: CSCA-TA, 8120 Woodmont Avenue, Bethesda, Maryland 20814-2797.

THE US ARMY'S CENTER FOR STRATEGY AND FORCE EVALUATION



# STOCHEM INVESTIGATION OF COSAGE SAMPLING (SICS)

1



## Background for STOCCEM investigation

- STOCCEM was developed as a stochastic version of CEM.
  - ✓ Use of deterministic CEM should be treated as an attempt to approximate the STOCCEM mean of many replications.
- Several STOCCEM investigations showed frequent and substantial divergence of deterministic CEM results from the STOCCEM mean. These results:
  - ✓ may imply that deterministic CEM and STOCCEM are not versions of the same model.
  - ✓ may imply that STOCCEM behavior is extremely nonlinear.
    - for a nonlinear mathematical function  $F(x)$ , expected value of  $F(x)$  diverges from  $F(\text{expected value of } x)$ .
  - ✓ may imply that deterministic CEM is a good approximation of STOCCEM *only if inputs for the two models are properly aligned.*

2

The Stochastic CEM (STOCCEM) theater combat simulation was developed by the US Army Concepts Analysis Agency (CAA) to be a stochastic version of the Concepts Evaluation Model IX (CEM IX) deterministic simulation. CEM IX uses only fixed inputs for the processes which are modeled stochastically in STOCCEM. The results of deterministic CEM are used as an expected simulation outcome in many CAA studies.

Several internal CAA studies (References 1, 2, 3) have found frequent and considerable differences between results of deterministic CEM and the mean STOCCEM result under a common scenario. The causes of these differences have not been fully determined, but may include one or more of the following:

- a. Although STOCCEM was intended to be a stochastic version of CEM IX, design differences may have caused some STOCCEM processes to differ substantially from, or be lacking in, a deterministic CEM counterpart. In such a case, two different models will likely produce radically different results.
- b. Based on mathematical principles, for a function  $F$ , in general, [expected value of  $F(x)$ ] is not identical to  $F(\text{expected value of } x)$ . This is always true only when  $F$  is linear. Substantial differences can arise when this is not the case.
- c. Results from STOCCEM and CEM under a common scenario can only be compared if the deterministic CEM inputs reflect the central tendency of the STOCCEM processes. If inputs are not aligned in this manner, significant divergence can occur.





## **Background (continued)**

- **Ardennes Campaign Simulation (ARCAS) study used STOCCEM.**
  - ✓ Variation in non-FEBA STOCCEM results (over replications) was often small.
  - ✓ Alternative ways of processing COSAGE samples in STOCCEM may increase variability.
- **A best way of using STOCCEM must be determined for the Kursk Operation Simulation and Validation (KOSAVE) study.**

3

STOCCEM was applied in the Ardennes Campaign Simulation Study (ARCAS) (Reference 4). Variation, over replications, in ARCAS STOCCEM measure of effectiveness (MOE) results often appeared to be unusually small for MOEs other than FEBA progress.

This variation can be increased by simply increasing the values of STOCCEM inputs regulating variation of selected statistical stochastic processes in that model. However, analysis of stochastic processes applied in STOCCEM suggested that variability in weapon effects may have been dampened to an unknown degree by the combat effects sampling methodology applied in STOCCEM. In ARCAS, a single combat effects sample was applied to all simulated engagements in a STOCCEM replication. Considering variation over engagements not only appears appropriate, but may increase variability in STOCCEM results sufficient to increase the likelihood that deterministic CEM results fall within the STOCCEM outcome range. Therefore, the SICS investigation focused on assessing variability resulting from alternative methods of STOCCEM use of combat samples which are generated by the Combat Sample Generator (COSAGE) simulation. No STOCCEM input parameter regulates variability over COSAGE samples.

Use of lessons learned in ARCAS to investigate and improve the use and interpretation of STOCCEM is needed, so that STOCCEM can be applied in the ongoing KOSAVE Study, which has objectives analogous to those in ARCAS.

THE US ARMY'S CENTER FOR STRATEGY AND FORCE EVALUATION

CAA  
CONCEPTS ANALYSIS AGENCY

## PURPOSE

- Investigate alternative methods of COSAGE sampling for use in the STOCCEM simulation.

4

The purpose of the SICS Research Analysis Activity (RAA) is to develop a preferred way to use COSAGE samples in general STOCCEM applications, including the KOSAVE Study.



### **Objectives for STOCCEM process improvement**

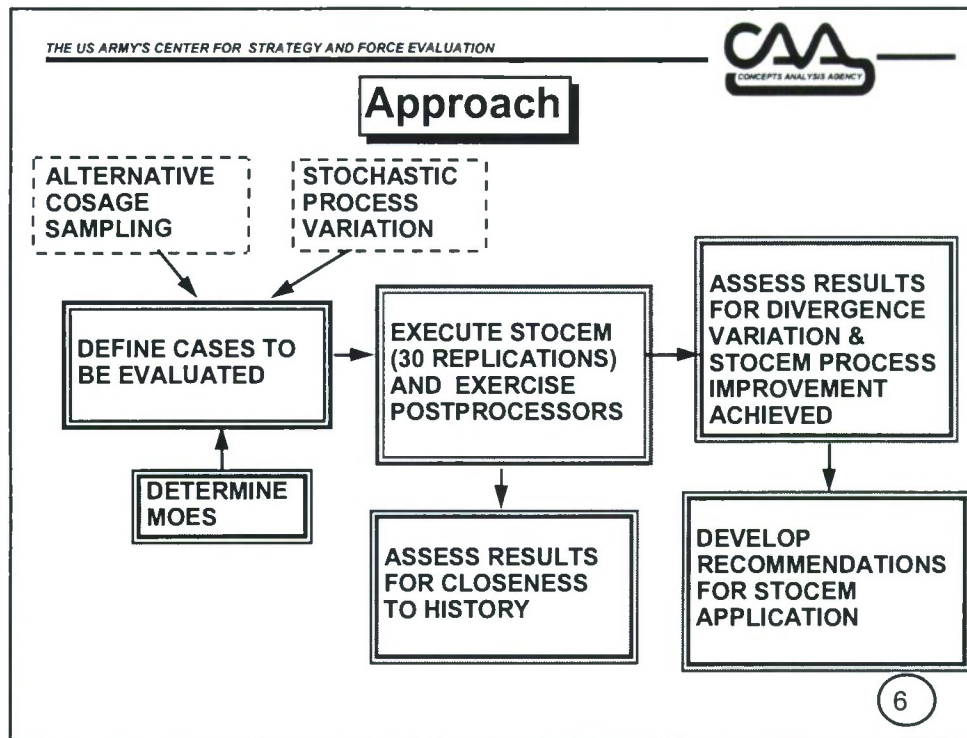
- **Reduce divergence of STOCCEM results from deterministic CEM results.**
- **Increase variability (over replications) of non-FEBA STOCCEM results.**
  - ✓ **Avoid increasing variation in FEBA results.**
- **Assess closeness of STOCCEM results to historical (ARCAS) results).**

5

Since ARCAS demonstrated a considerable, and sufficient, variability in STOCCEM FEBA progress results over replications, an increase in variability, over replications is deemed desirable only in non-FEBA MOE results, i.e., those MOEs which do not measure FEBA progress.

A secondary objective of SICS is to assess the closeness of STOCCEM results to historical results in the base STOCCEM case, and when alternative COSAGE sampling methodology is used.





Using the ARCAS scenario as a base, the SICS analytic steps are to:

- a. Define alternatives to the present method of COSAGE sampling which are directed toward achieving the SICS objectives.
- b. Define variations in the use of stochastic processes in STOCEM which are directed toward achieving the SICS objectives.
- c. Define the STOCEM cases to be evaluated. These are defined in terms of the COSAGE sampling alternatives and stochastic process variations which are applied in each case.
- d. Define the STOCEM MOEs to be evaluated.
- e. Execute each STOCEM case for 30 replications and exercise postprocessors to compute the mean STOCEM MOEs, as well as the variability over replications. Exercise deterministic CEM once.
- f. Assess STOCEM case results relative to divergence from deterministic CEM results and to variability over replications.
- g. Using the SICS objectives, select a preferred case and develop recommendations for improved use of COSAGE sampling in STOCEM.
- h. Assess STOCEM results for closeness to historical Ardennes Campaign results when alternative COSAGE sampling is applied.



## Measuring STOCem divergence from deterministic CEM

- **Assess frequency that deterministic CEM MOE values are outside following STOCem limit bands:**
  - ✓ 25th percentile & 75th percentile (over 30 replications)
  - ✓ minimum & maximum (over 30 replications)
  - ✓ 99 % confidence limits (around STOCem mean)
  - ✓ STOCem mean  $\pm$  15% (of STOCem mean)
- **STOCem process improvement objective:**
  - ✓ Reduce frequency of deterministic CEM MOEs outside STOCem limit bands

7

Divergence of a deterministic CEM MOE from the corresponding STOCem mean MOE is assessed in terms of statistical as well as absolute divergence. Statistical divergence is defined as the frequency with which deterministic CEM MOE values lie outside specified statistical bounds surrounding the corresponding STOCem mean, e.g., percentiles or confidence limits. Statistical divergence is an indicator of divergence relative to the full distribution of STOCem rather than just the mean value. For the range of MOEs examined, divergence is computed as the fraction of the deterministic MOEs with values outside the four STOCem bounds/limits described in this chart. All STOCem confidence limits are based on the Student *t*-distribution being applied to sample means.



## Measuring variation (over replications) in STOCCEM results

- **Compute standard deviation of results (over 30 replications)**
  - ✓ Comparison of standard deviations is equivalent to comparison of confidence limit bands.
- **STOCCEM process improvement objective:**
  - ✓ Increase the standard deviation of non-FEBA STOCCEM MOEs.

8

Variability of STOCCEM MOE results over replications is measured by the standard deviation of the sample of 30 replication results for an MOE. The width of a confidence interval is directly proportional to the standard deviation of the underlying sample distribution.

Treating STOCCEM application as a process, one objective of SICS is to modify the STOCCEM process to increase the sample standard deviation of MOE results for MOEs other than FEBA progress.





## Assessing closeness of STOCEM results to history

- Use historical Ardennes campaign results derived from the Ardennes Campaign Simulation Data Base (ACSDB) as a baseline.
  - ✓ Quantified in ARCAS Study
  - ✓ Compare only STOCEM mean results for selected cases.
    - (1) Base STOCEM case
    - (2) Deterministic CEM case
    - (3) Cases using alternative COSAGE sampling

The ARCAS Study (Reference 4) and the SEACA RAA both compared STOCEM results from a simulation of the 1944-45 Ardennes Campaign with historical results derived from the ACSDB. Those historical results served as an historical baseline for comparison with SICS STOCEM results simulating the same campaign.

Historical comparison in SICS was done only with the base STOCEM case, as applied in the SEACA RAA, the deterministic CEM cases, and the STOCEM variants (described in Chart 12) which apply alternative COSAGE sampling (relative to the method used in the base case). All STOCEM cases consisted of 60 STOCEM replications.



## **Alternative COSAGE sampling**

- **Current STOCCEM practice is to draw a COSAGE sample for every subsector battle.**
  - ✓ In this case, mean theater weapon effectiveness in each replication reflects the average COSAGE sample.
- **Alternative COSAGE sampling is defined as drawing only one sample per replication for each posture.**
  - ✓ This case explicitly models variation in mean theater weapon effectiveness over replications.

10

In current STOCCEM practice, in each replication, a uniform random draw of a COSAGE combat sample is made for each subsector engagement in each STOCCEM division cycle (usually 12 simulation hours). This draw is from the replications (samples) of the COSAGE-generated killer/victim scoreboards associated with the combat posture (e.g., Red attack vs Blue hasty defense) of that engagement. Since a STOCCEM division cycle has many subsector engagements, the law of large numbers suggests that the the full theater effect in a STOCCEM theater cycle (usually 4 simulation days) reflects the central tendency of the distribution of COSAGE samples. As a result, the mean STOCCEM theater weapon effectiveness in each STOCCEM replication will always approximately reflect the average COSAGE sample.

Define alternative COSAGE sampling in STOCCEM as the random drawing of a COSAGE replication for each combat posture at the start of a replication, and use of that fixed set of samples for all (subsector) engagements of all division cycles in that replication. With this alternative sampling, the mean STOCCEM theater weapon effectiveness in each replication approximately reflects the specific COSAGE samples drawn at the start of the replication. Variation in mean STOCCEM theater weapon effectiveness over replications is explicitly modeled in this process. Alternative COSAGE sampling, as defined in this manner, is also a direct generalization of COSAGE sampling as applied in deterministic CEM, which uses a single fixed set of average combat samples throughout the simulation.



## **Stochastic processes used in STOCCEM**

- **FEBA move rate produced by a tactical situation.**
- **Decision thresholds for allocation & commitment.**
- **Threshold for hasty (vs prepared) defense posture.**
- **Fraction damaged vehicles destroyed & abandoned.**
- **Fraction personnel casualties wounded, hospitalized, evacuated.**
- **Weapon effectiveness (combat samples)**

11

This list summarizes the stochastic processes that can be simulated in STOCCEM. A more detailed description of their definition and operation can be found in STOCCEM documentation (References 1, 2, 3).





## Definition of STOCCEM cases to be evaluated

CASE ID	DETERMINISTIC STOCCEM PROCESSES	COMBAT SAMPLING
B0 (base)	None	per subsector battle
B2	FEBA movement	per subsector battle
B2	FEBA movement, decision thresholds	per subsector battle
B2	All except combat samples	per subsector battle
C2	None	1 per theater/replication
C2	FEBA movement	1 per theater/replication
C2	FEBA movement, decision thresholds	1 per theater/replication
C4	All except combat samples	1 per theater/replication
DET CEM	All	1 for theater
DCS	Combat samples (average sample is used)	1 for theater

12

Case B0 denotes the standard (current) way of running STOCCEM, in which all of the stochastic processes described in Chart 11 are activated. B1 is a variant case which makes only the FEBA movement process deterministic (all other processes are as in B0). B2 is a variant which makes only the FEBA movement process *and* the decision threshold process both deterministic. The B4 variant makes all of the usual (in B0) STOCCEM stochastic processes deterministic except for COSAGE sampling, which is stochastic. All of the C cases (C0, C1, C2, and C4) are exactly analogous to the B cases (B0, B1, B2, and B4) except that the C cases all select a combat sample only once in each replication. This differs from the B cases which select a combat sample for each subsector engagement/battle in a replication. The DET CEM case denotes deterministic CEM.

The Deterministic COSAGE with Stochastic (DCS) case denotes the case in which all the usual (in B0) STOCCEM stochastic processes are activated except for COSAGE sampling, which is deterministic. In the DCS case, the fixed COSAGE sample used in STOCCEM is exactly the same as that used in the DET CEM case. This sample is a composite average over all samples.

THE US ARMY'S CENTER FOR STRATEGY AND FORCE EVALUATION



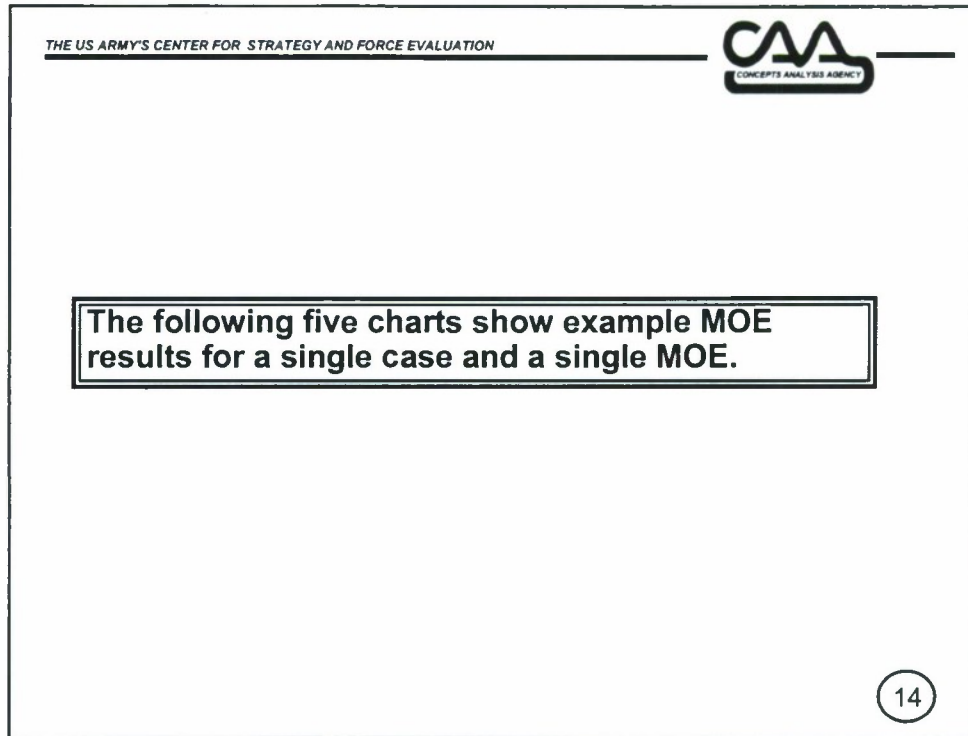
### MOEs examined

- 14 theaterwide MOEs for each theater cycle (TC) include:
  - ✓ mean FEBA change & average net FEBA change (2 MOEs)
  - ✓ Blue ammo spent/TC & cumulative Blue ammo spent (2 MOEs)
  - ✓ Red ammo spent/TC & cumulative Red ammo spent (2 MOEs)
  - ✓ Blue & Red cumulative permanent tank losses (2 MOEs)
  - ✓ Blue permanent personnel losses (by TC & cumulative 2 MOEs)
  - ✓ % Blue force in static posture
  - ✓ % Blue force in reserve posture
  - ✓ % Blue force in attack posture
  - ✓ % Blue force being attacked by Red
- FEBA position MOEs measure cumulative FEBA progress (since D-day) on each STOCES avenue of advance.

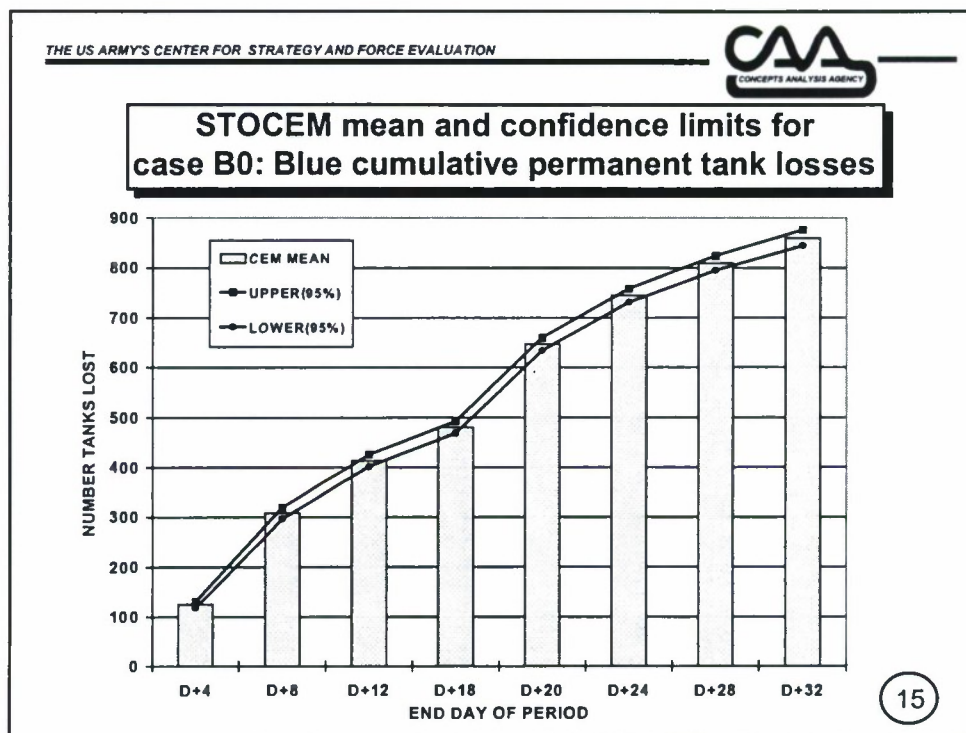
13

The CEM/STOCES MOEs examined in SICS are divided into two groups. The first group, denoted in the charts as 14 theaterwide MOEs, comprises 14 MOEs from the Automated Data Display of CEM Outputs (ADDCOP) files, which are automatically compiled by the CEM Report Generator during each model execution. These output reports are described more fully in CEM documentation (Reference 5). Each of the 14 theaterwide MOEs is computed for the entire Red (German) or Blue (US/UK) force for each of the eight 4-day theater cycles in the Ardennes Campaign. FEBA advance and change are arithmetically averaged over all CEM avenues of advance defined in the theater. Ammunition spent is measured in tons. Permanent personnel losses include killed, captured, missing, and patients evacuated from theater. Permanent tank losses include only killed and abandoned systems. Over the 32-day campaign, the entire group comprises exactly 112 MOE values generated by deterministic CEM, and also during each STOCES replication. The arithmetic mean value over all 30 replications of a STOCES case is the primary STOCES MOE used for comparison with the corresponding deterministic CEM result.

The group denoted as FEBA position MOEs is defined as the FEBA progress (by the Germans) along each of the 21 active avenues of advance used by forces in the simulation. Their specific configuration is described in ARCAS documentation. Over the 8 theater cycles of the full campaign, this group comprises exactly 168 MOE values generated by deterministic CEM, and during each STOCES replication. The arithmetic mean value over all 30 STOCES replications is used for comparison with deterministic CEM.

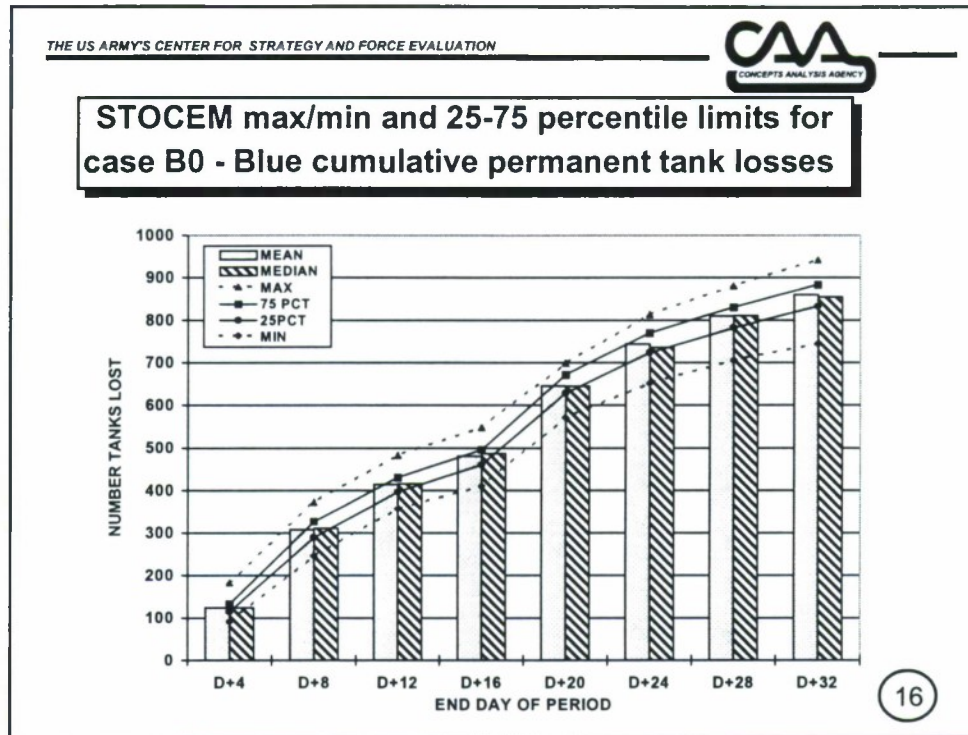


The following five charts show example “bottom level” graphic results used to build the comparative results described later. Each example is for a case defined in Chart 12 and and MOE described in Chart 13.

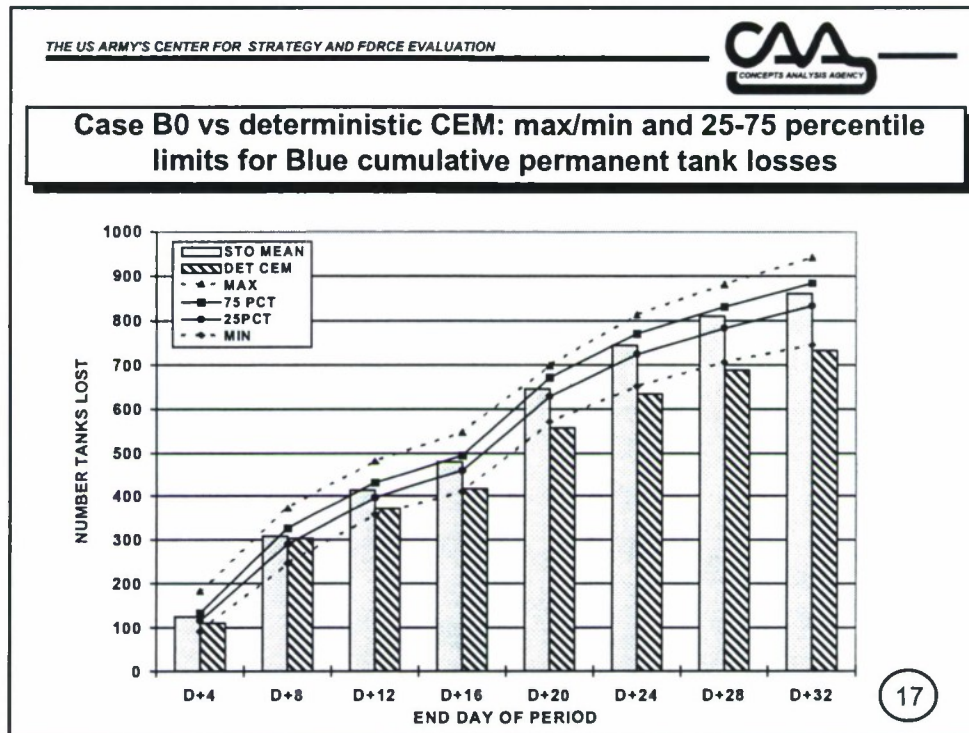


This example shows the STOCEM means, and 95 percent confidence limits, for the sampling distribution of the mean, for example results from case B0 for each 4-day CEM theater cycle. STOCEM means are arithmetic averages over 30 replications in the case (B0). Confidence limits around each STOCEM mean are computed using the Student  $t$ -distribution. The "Blue" force denotes the US/UK in ARCAS results.





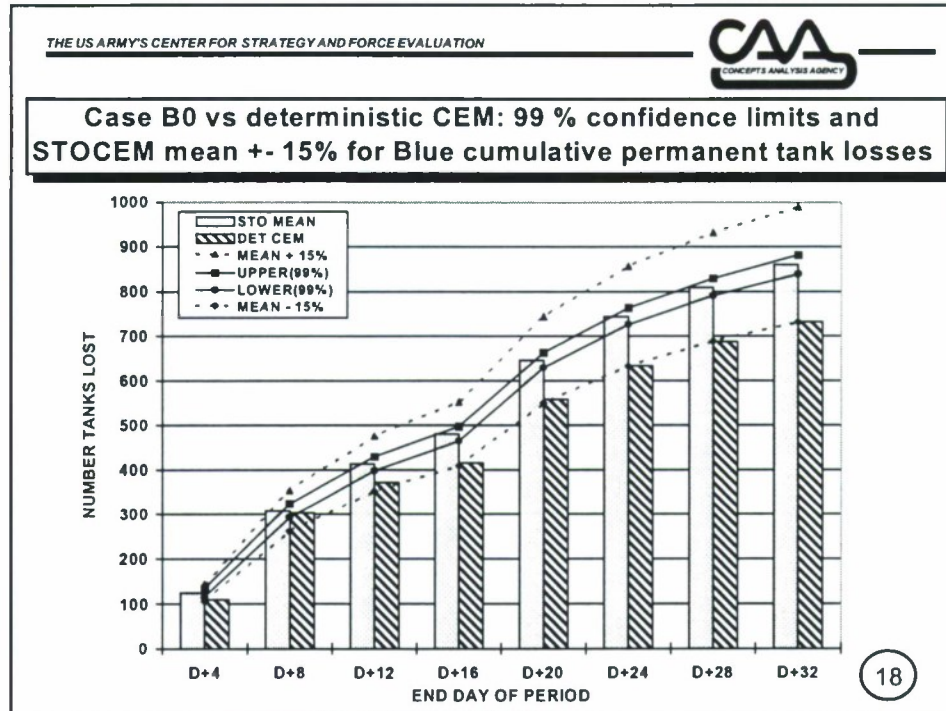
This chart shows attributes of the statistical distribution of STOCem results for the case depicted. The minimum/maximum bounds, and the 25th/75th percentile bounds are two of the four types of STOCem limit bands noted in Chart 7.



This chart shows deterministic CEM results for the cumulative Blue (US/UK) permanent tank loss MOE for each 4-day theater cycle, along with:

- The corresponding case B0 STOCCEM mean for the MOE (over 30 replications).
- The 25th percentile limit (labeled 25 PCT) and the 75th percentile limit (labeled 75 PCT) for the case B0 STOCCEM MOE (over 30 replications).
- The minimum bound and maximum bound for the case B0 STOCCEM MOE (over 30 replications).

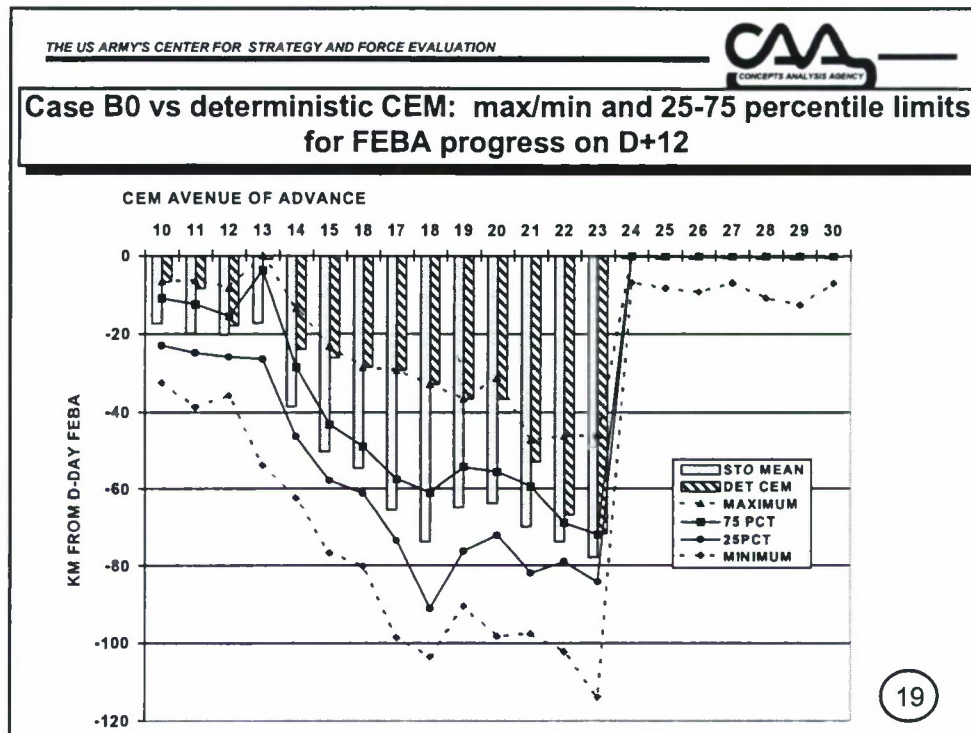
This chart graphically shows example comparative results for two of the four STOCCEM limit bands noted in Chart 7.



This chart shows deterministic CEM results for the cumulative permanent Blue (US/UK) tank losses MOE for each 4-day theater cycle, along with:

- The corresponding case B0 STOCEM mean for the MOE (over 30 replications).
- The 99 percent confidence limit band around the case B0 STOCEM mean (over 30 replications).
- The values of [85 percent of the B0 STOCEM mean] and [115 percent of the B0 STOCEM mean] .

This chart graphically shows example comparative results for two of the four STOCEM limit bands defined in Chart 7.



The deterministic CEM FEBA cumulative progress on D+12 is plotted in this chart for each of the 21 STOCCEM avenues of advance, which are indexed left-to-right corresponding to a north-to-south geographic order. The magnitude of the FEBA progress is plotted for each avenue of advance. The D-day (Day 1) position is at the 0 ordinate, and a negative "km from D-day FEBA" corresponds to a German advance. This linearized representation emulates a quasi-geography for the battle with relative positions along the STOCCEM avenues of advance, which are here represented as parallel straight lines. The orientation is from an stylized aerial perspective facing east from above US/UK lines.

a. The points on the dashed line graphs show the maximum and minimum (westernmost and easternmost) case B0 STOCCEM FEBA progress since D-day over the 30 replications.

b. Points on the thin solid line graphs in the figure (which are labeled 25 PCT and 75 PCT) show the 25th percentile limits and the 75th percentile limits for the distribution, over 30 replications, of case B0 STOCCEM FEBA progress since D-day, in kilometers (km).

This chart graphically shows example comparative results for two of the four STOCCEM limit bands noted in Chart 7.

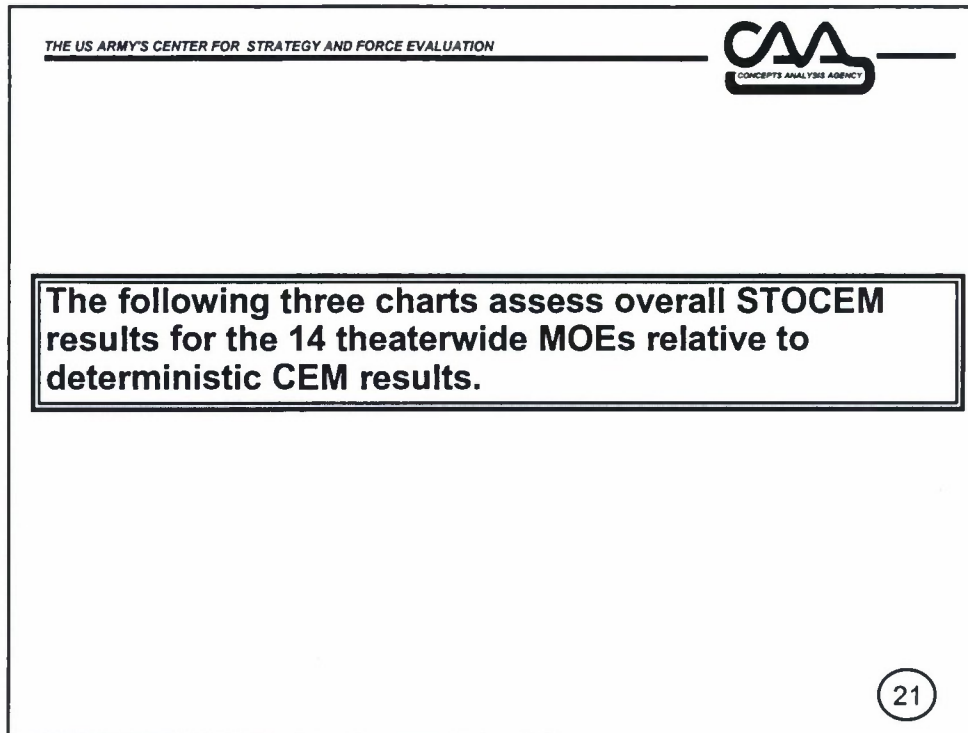


### Bias in STOCCEM FEBA movement process

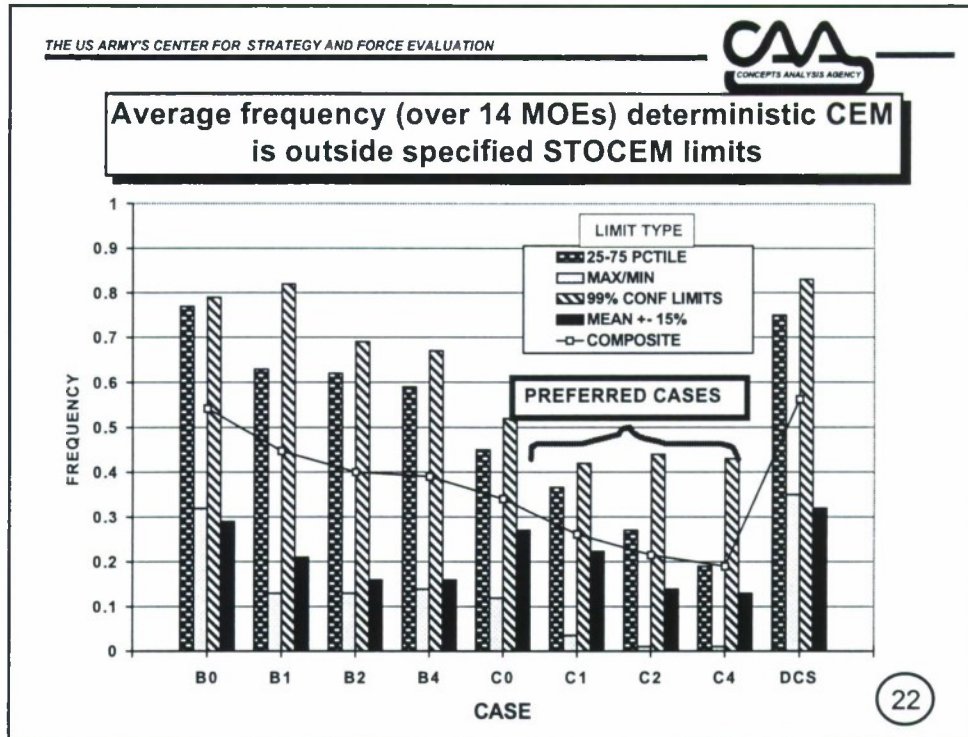
- The STOCCEM FEBA move rate process is probably biased (having a different mean) relative to FEBA move rates generated in deterministic CEM.
  - ✓ Central tendency of STOCCEM "rate of advance" tables was not calibrated to deterministic CEM values.
  - ✓ All other STOCCEM processes are calibrated to same mean as deterministic CEM values.
  - ✓ FEBA movement bias may increase divergence from deterministic CEM FEBA movement.
- Base case B0 and case C0 are *a priori* not preferred.
  - ✓ Results are generated so that any bias may be exhibited.

20

Conversion of a deterministic simulation to a stochastic form is conventionally done by defining each stochastic process as a statistical distribution with a central tendency (usually the mean) equal to either an input of the deterministic model, or to an MOE generated in the deterministic model. This calibration process does apply to all of the stochastic processes of STOCCEM except the FEBA move rate process. Discussions with the STOCCEM designer indicate that the stochastic FEBA movement process in STOCCEM was based on updated historical move rate data and that no attempt was made to calibrate it to the move rate process used in deterministic CEM. The STOCCEM FEBA movement is therefore probably biased to an unknown degree, relative to FEBA movement in deterministic CEM. Only SICS cases B0 and C0 activate the STOCCEM FEBA move rate stochastic process. Because of the above bias, these two cases are *a priori* not suitable for comparison with deterministic CEM to assess differences due to alternative COSAGE sampling. However, comparisons for these cases are included herein, so that any bias may be quantitatively demonstrated.



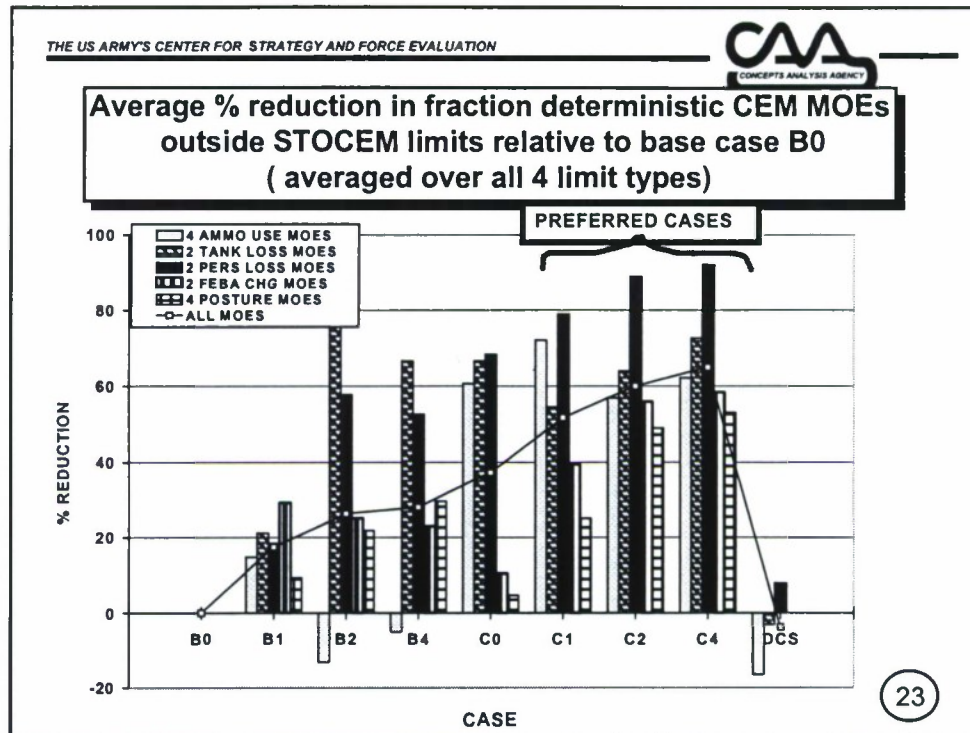
The chart is self-explanatory.



This chart shows and compares nine clusters of case results, corresponding to, and labeled with, the names of the nine STOCEM cases studied. B0 denotes the base STOCEM case, B1 through B4 denote the variations of the base, in which increasing numbers of stochastic STOCEM processes are made deterministic. The cases are described/defined in Chart 12.

Each bar of each case result cluster shows the fraction of the 14 (MOEs) x 8 (theater cycles) = 112 deterministic CEM theaterwide MOE results for that case which are outside of the STOCEM limits (from Chart 7) associated with that bar. For example, the leftmost (striped) bar of each cluster denotes the 25-75 percentile limits of the case (which by definition bound 50 percent of the STOCEM results for case). Looking at B0, we see that the ordinate, 77 percent, describes the fraction of the 112 deterministic CEM MOEs which are either less than the 25 percentile limit or greater than the 75 percentile limit for the case. Each bar type can be compared across case clusters. A low value indicates less divergence relative to the indicated STOCEM limits. Each small white square in a case cluster is the arithmetic average of the ordinates for the bars in that cluster. These squares are labeled as "composite" and are connected by a line solely for ease of viewing.

The three indicated/bracketed cases, C1, C2, and C4, show the least divergence and are hereafter (in this briefing) labeled, and denoted, as the *preferred cases*. These use the alternative COSAGE sampling and exclude (biased) stochastic treatment of the FEBA move rate process.



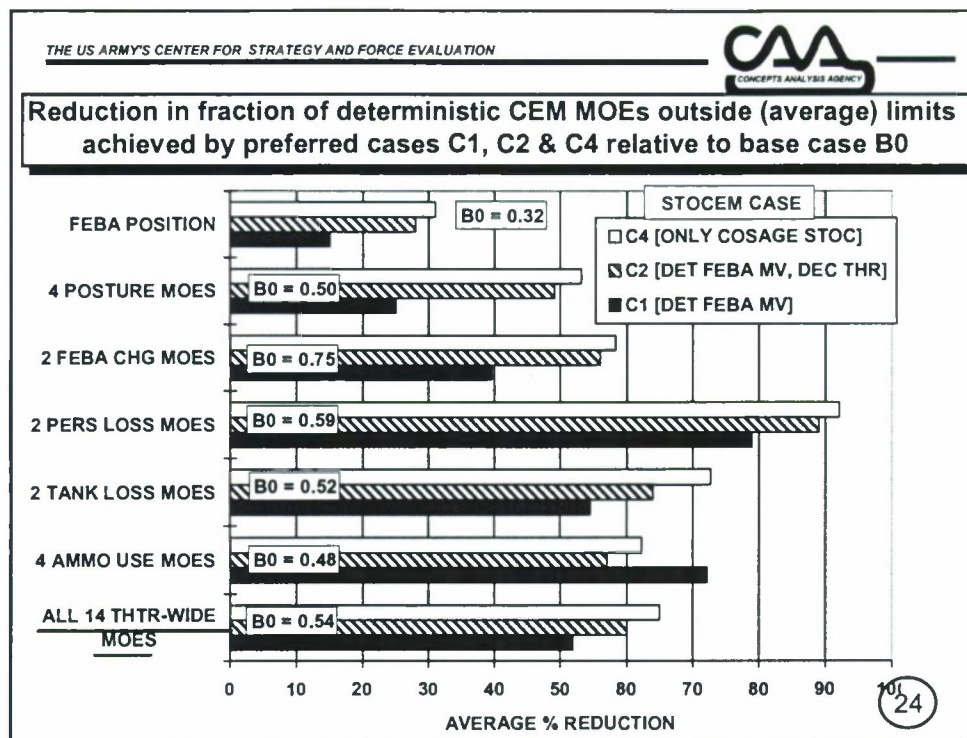
This chart also shows and compares results for nine STOCCEM cases studied. However each bar of each case result cluster here is associated with an MOE type rather than a STOCCEM limit type. The ordinate of each bar shows the average reduction, relative to the base case B0, in the fraction of deterministic CEM MOE results, of the specified MOE type for that case, which are outside of STOCCEM limits. Fraction outside STOCCEM limits is defined as the arithmetic average of the fractions (outside limits) for each of the four limit types. The "ALL MOES" value in a case cluster is the arithmetic average of the ordinates for the four bars in the cluster and reflects the relative change (in terms of reduction) relative to case B0 in the case values for the composite limit in the previous chart. For example, the composite fraction outside limits for B0 in the previous chart was .55 for case B0 and .45 for case B1. For case B1, this reflects a reduction of  $10/55 = .18$  (= 18 percent) relative to B0 and this is the B1 ordinate for "ALL MOES" in this chart.

In this chart, a high value indicates less divergence relative to the associated STOCCEM limits.

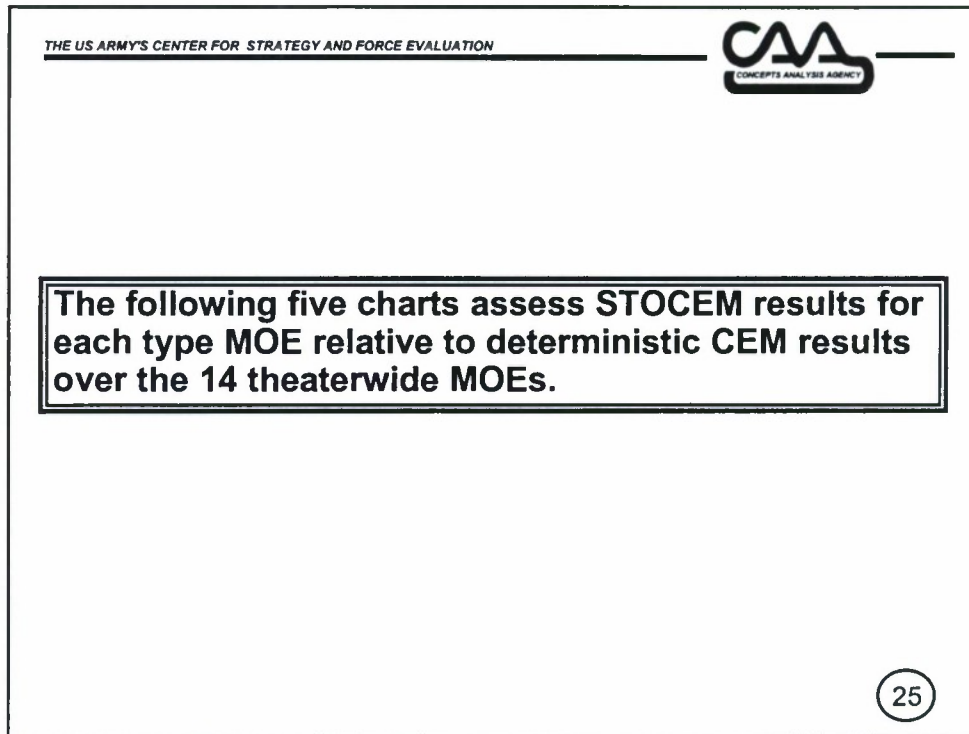
The three indicated preferred cases, C1, C2, and C4, show the least divergence. These use the alternative COSAGE sampling and exclude stochastic treatment of the FEBA move rate. The DCS case, which has everything stochastic except the fixed COSAGE sample, diverges even more than the base case.

The biggest reductions are associated with the personnel loss MOEs. The smallest reductions are associated with the "percent in posture" MOEs.





This chart compares average percent reductions in the fraction outside STOCEM limits, relative to the base case B0, by MOE type, for only the designated (in Charts 22 and 23) preferred cases C1, C2, and C4. These percent reduction values are from the previously shown chart and reflect averages over all four STOCEM limit types. The fraction outside limits for the base case B0 is shown in the inset white box for each MOE type. Except for posture and FEBA MOEs, there are only small differences between percent reduction achieved by the three preferred cases. These preferred cases, over all 14 theaterwide MOEs, yield an average 50 to 65 percent reduction in divergence relative to the base case B0. However for the FEBA position MOEs, these three preferred cases, using alternative COSAGE sampling, yield only a 15 to 35 percent average reduction in divergence measures.



The chart is self-explanatory.

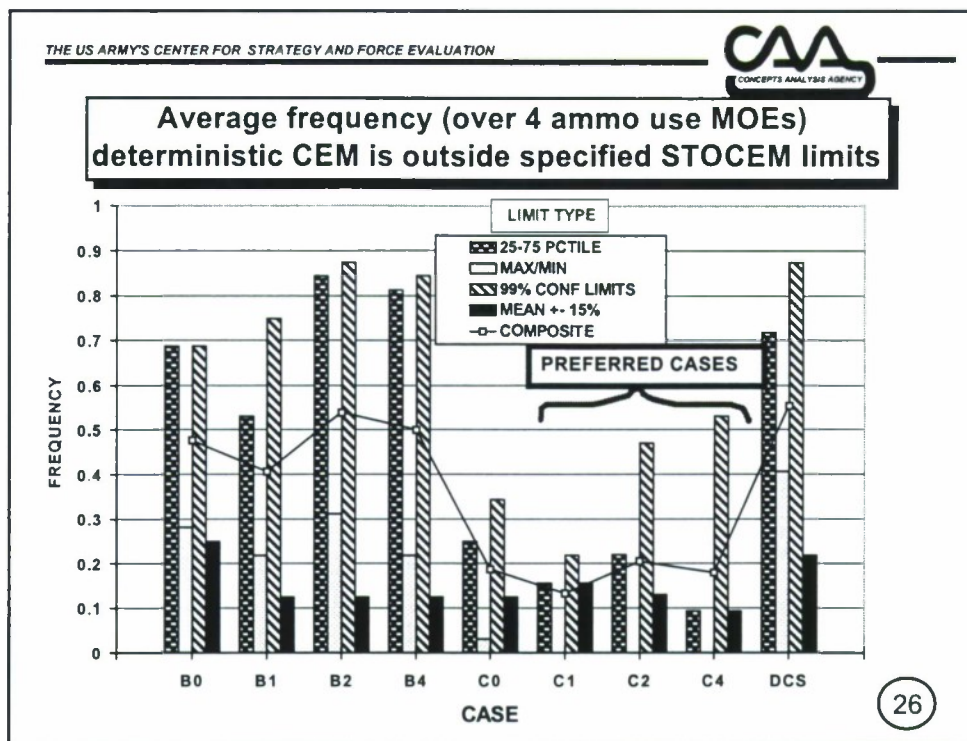
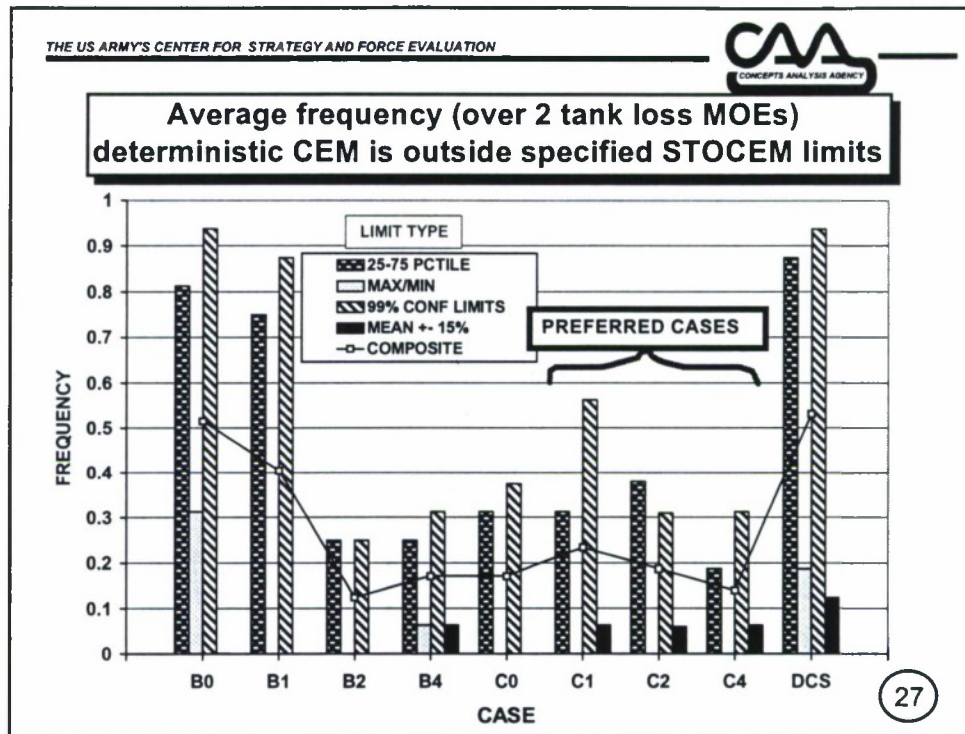


Chart 22 showed the average fraction of all deterministic CEM theaterwide MOEs outside specified STOCem limits for each case. This, and the next four charts, show the average fraction outside limits only for the MOEs of a specific type (chosen from the 14 theaterwide MOEs described in Chart 13). This chart shows the average fraction outside limits only for the four ammo use MOEs considered. (This comprises 32 deterministic CEM MOEs.) The nine case clusters are in exactly the same format as shown in Chart 22.

Each bar of each case result cluster here shows the fraction of the  $4 \times 8 = 32$  deterministic CEM ammo use MOE results for that case which are outside of the STOCem limits associated with that bar. The leftmost bar of each cluster denotes the 25-75 percentile limits of the case. Looking at the B0 case cluster, we see that the ordinate 69 percent describes the fraction of the 32 deterministic CEM MOEs which are either less than the 25 percentile limit or greater than the 75 percentile limit for the case. Each bar type can be compared across case clusters. A low value indicates less divergence relative to the indicated STOCem limits. In this, and in the following four charts, the white square in each case cluster, denoted as composite, is the arithmetic average of the ordinates for the four bars in the case cluster. These white squares are connected by a line solely for ease of viewing.

The three indicated preferred cases (showing least divergence in Charts 22 and 23) also show the least divergence for this MOE if case C0 is excluded. These preferred cases all use the alternative COSAGE sampling.

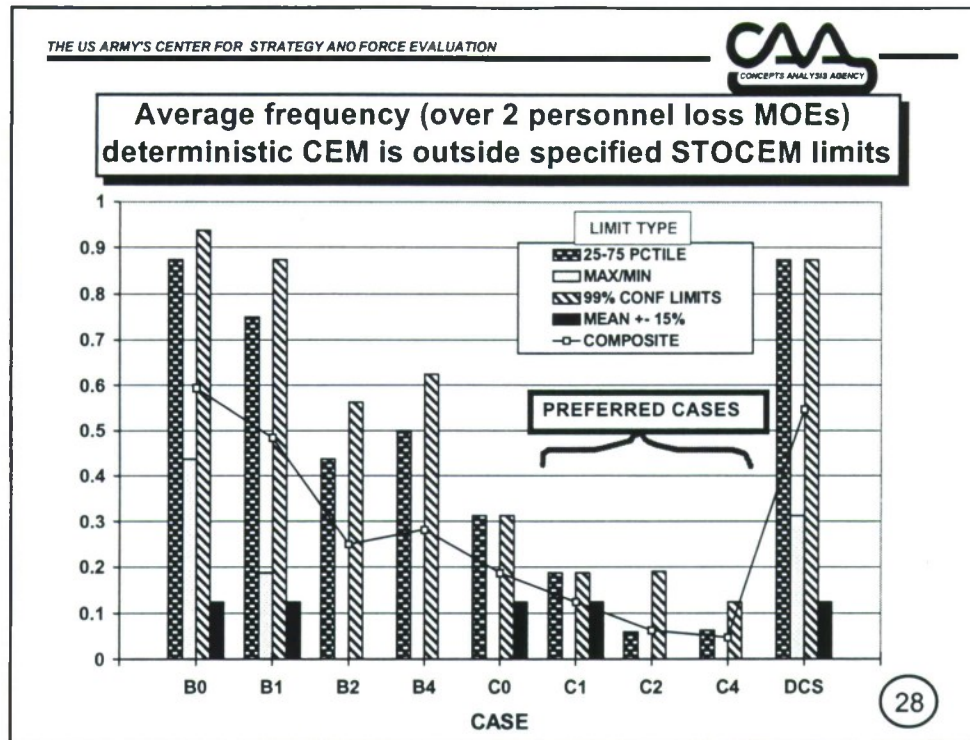


Each bar of each case result cluster here shows the fraction of the  $2 \times 8 = 16$  deterministic CEM tank loss theaterwide MOE results for that case which are outside of the STOCCEM limits associated with that bar. Each bar type can be compared across case clusters. A low value indicates less divergence relative to the indicated STOCCEM limits.

The three indicated preferred cases, C1, C2, and C4, are those selected previously, in Charts 22 and 23, as displaying least overall divergence. For this MOE type, divergence of the preferred cases C2 and C4 is not very different from their B2 and B4 counterparts.

The DCS case, which activates every STOCCEM stochastic process except the fixed COSAGE sample, again shows substantial divergence.

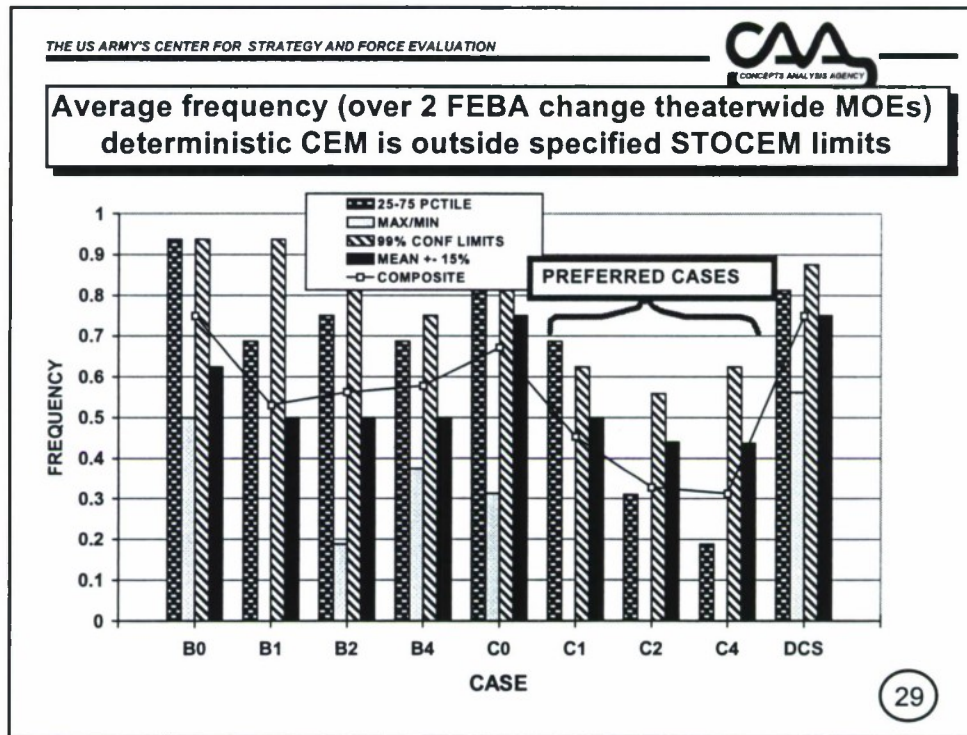




Each bar of each case result cluster here shows the fraction of the  $2 \times 8 = 16$  deterministic CEM personnel loss MOE results for that case which are outside of the STOCCEM limits associated with that bar. A low value indicates less divergence relative to the indicated STOCCEM limits. The white square in each cluster, denoted as composite, is the arithmetic average of the ordinates for the four bars in the case cluster.

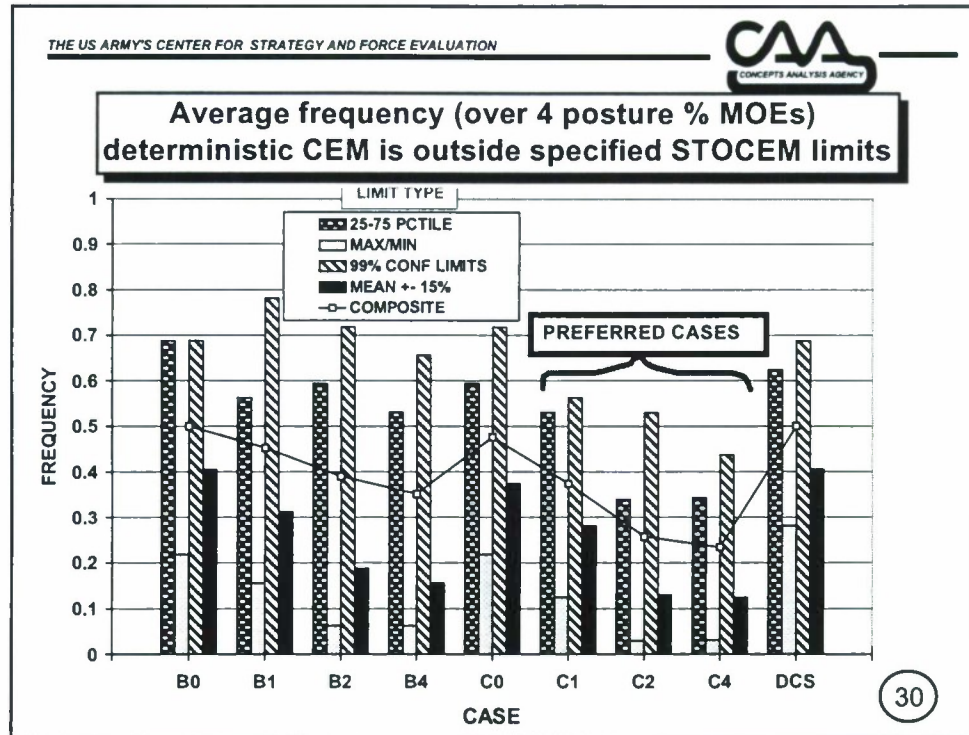
The three indicated preferred cases, C1, C2, and C4, from Charts 22 and 23, also show the least divergence for this MOE.

The DCS case, which has everything stochastic except the fixed COSAGE sample, again shows substantial divergence.



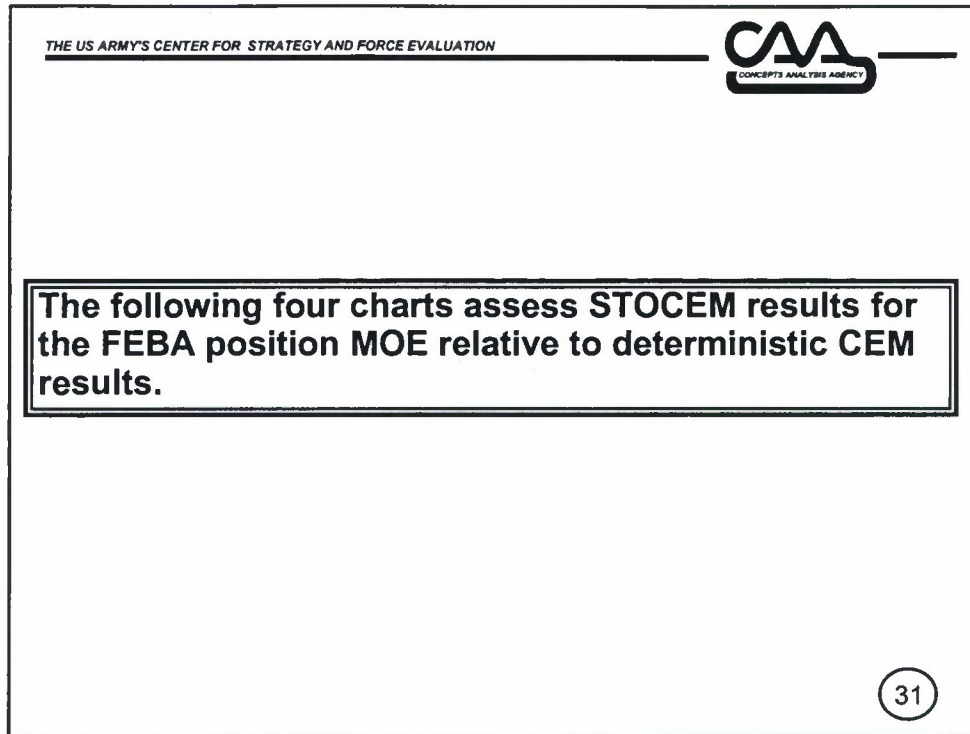
Each bar of each case result cluster here shows the fraction of the  $2 \times 8 = 16$  deterministic CEM “FEBA change” theaterwide MOE results for that case which are outside of the STOCCEM limits associated with that bar. A low value indicates less divergence relative to the indicated STOCCEM limits. The white square in each case cluster, denoted as composite, is the arithmetic average of the ordinates for the four bars in the cluster.

The three indicated preferred cases, C1, C2, and C4, displaying least divergence in Charts 22 and 23, also show least divergence for this MOE type.



Each bar of each case result cluster here shows the fraction of the  $4 \times 8 = 32$  deterministic CEM results for the “percent force in posture” theaterwide MOEs which are outside of the STOCCEM limits for each STOCCEM case studied. A low value indicates less divergence relative to the indicated STOCCEM limits. The white square in each case cluster, denoted as composite, is the arithmetic average of the ordinates for the four bars in the cluster.

The three indicated preferred cases, C1, C2, and C4, displaying least divergence in Charts 22 and 23, also show least divergence for this MOE type, but they show only slightly less divergence than most other cases.



This chart is self-explanatory.



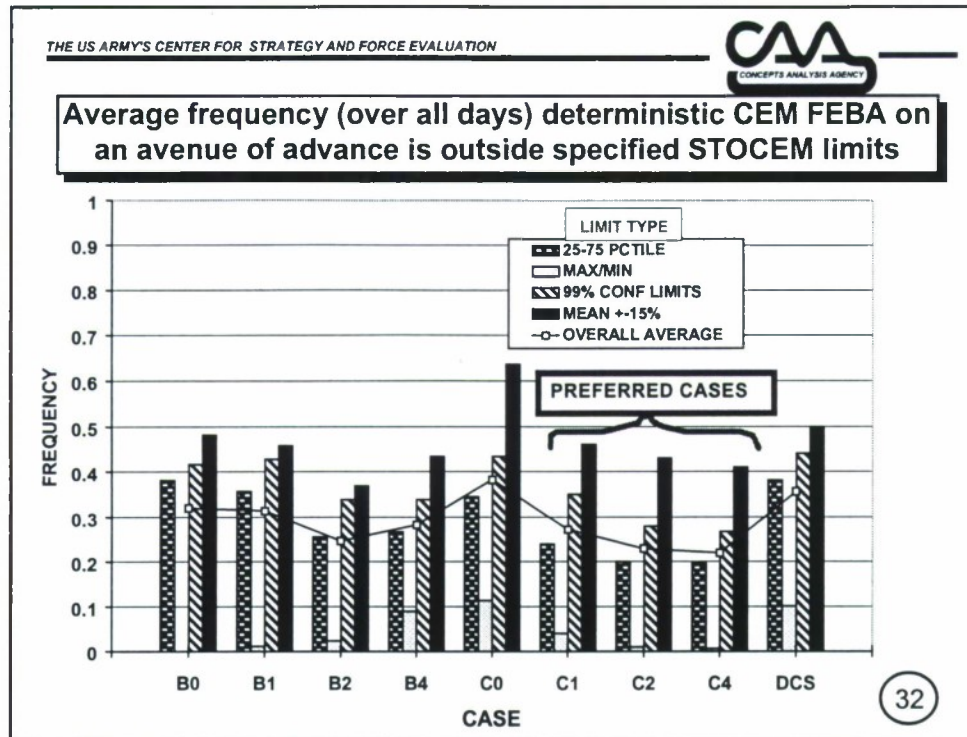
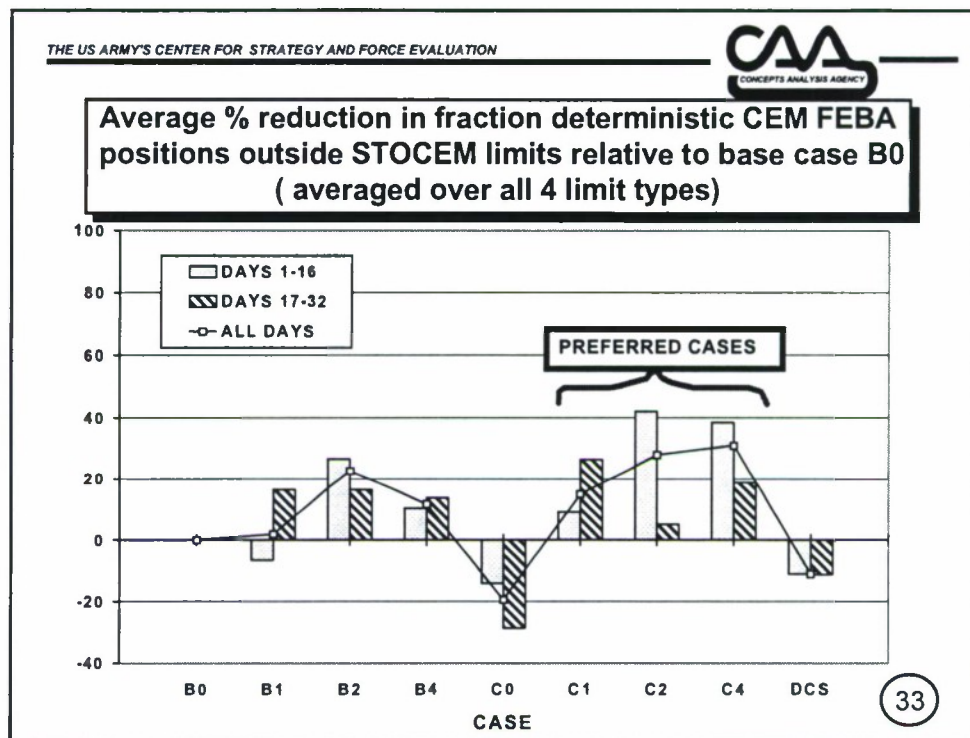


Chart 22 showed the average fraction of all 112 deterministic CEM theaterwide MOEs outside specified STOCCEM limits for each case. This chart shows the average fraction outside limits only for the FEBA position MOE described in Chart 13. It shows the average fraction of deterministic CEM FEBA position MOE results lying outside each of the specified STOCCEM limits over all of the 21 avenues of advance in the theater over all eight 4-day theater cycles. This chart has the same format as was used in Chart 22. Only the underlying MOEs are different.

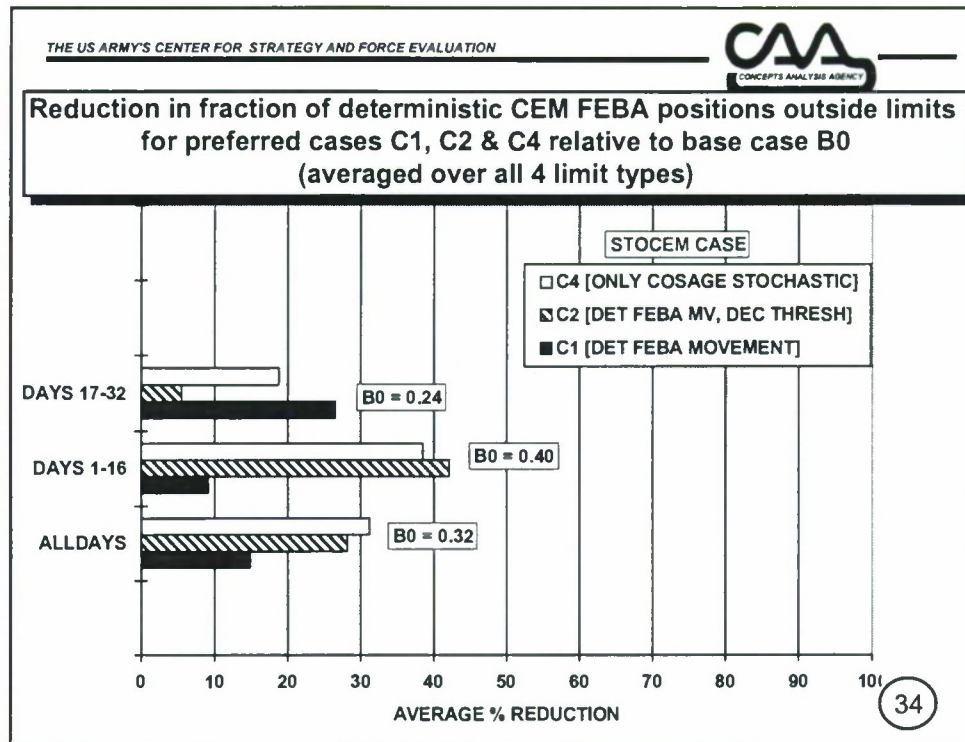
Each bar of each case result cluster here shows the fraction of the  $21 \times 8 = 168$  deterministic CEM FEBA position MOE results for that case which are outside of the STOCCEM limits associated with that bar. Again, the leftmost bar of each case cluster denotes the 25-75 percentile limits of the case. Each bar type can be compared across case clusters. A low value indicates less divergence relative to the indicated STOCCEM limits. The white square in each cluster, denoted as "overall average," is the arithmetic average of the ordinates for the four bars in the case cluster. These white squares are connected by a line solely for ease of viewing.

The three indicated preferred cases, C1, C2, and C4, are those which showed least overall divergence in Chart 22. These use the alternative COSAGE sampling and exclude (biased) stochastic treatment of the FEBA movement. For this MOE, these cases exhibit only slightly less divergence, on average, than their counterpart cases B1, B2, and B4



This chart also compares case results over the nine STOCCEM cases studied, but each bar of each case result cluster here is associated with a scenario timeframe rather than with a STOCCEM limit type. Specifically, the ordinate of each bar shows the average percent reduction, relative to the base case B0, in the fraction of the deterministic CEM FEBA position results for that case which are outside of STOCCEM limits for the indicated scenario timeframe. Fraction outside STOCCEM limits is defined here as the arithmetic average of the fractions (outside limits) for each of the four limit types. The white squares labeled “ALL DAYS,” and connected by a line, reflect the relative change (in terms of reduction) relative to case B0 over the entire scenario. For example, the overall average fraction outside limits for B0 in the previous chart was .32 for case B0 and .25 for case B2. For case B2 this reflects a reduction of  $.07/.32 = .21$  relative to B0 and this is the B1 ordinate for “ALL DAYS” in this chart. In this chart, a high value indicates less divergence.

The three indicated preferred cases are those selected previously, showing least divergence in Charts 22 and 23. These use the alternative COSAGE sampling and exclude (biased) stochastic treatment of the FEBA move rate. Here, they achieve somewhat greater reduction in average divergence than their B1, B2, and B4 counterparts. The DCS case, which activates all STOCCEM stochastic processes except the fixed COSAGE sample, shows a negative divergence reduction, indicating it has more divergence than the base B0 case. This is also true of case C0, and probably is substantially due to the bias in the FEBA move rate stochastic process, which is active only in cases B0, C0, and DCS.



This chart compares average percent reductions in fraction outside STOCES limits relative to the base case B0 for only the preferred cases C1, C2, and C4. These percent reduction values are from the previously shown chart. The fraction outside limits for the base case B0 is also shown in the inset white box for each timeframe. Percent reduction is arithmetically averaged over all four STOCES limit types. These preferred cases, over all 168 FEBA position MOEs, yield an average 15 to 35 percent reduction in divergence relative to the base case B0 over the entire scenario.

THE US ARMY'S CENTER FOR STRATEGY AND FORCE EVALUATION



## Observations on assessment of STOCCEM divergence from deterministic CEM

- **14 Theaterwide MOEs**
  - ✓ Best cases use alternative COSAGE sampling & have average 50 - 65 percent reduction in divergence measures.
  - ✓ Best case (65 percent reduction) C4 has only one activated stochastic process (COSAGE sampling)
  - ✓ Case C1 (50 percent reduction) has all stochastic processes activated except FEBA movement.
- **FEBA position MOEs**
  - ✓ Best cases use alternative COSAGE sampling & have average 15 - 35 percent reduction in divergence measures.
  - ✓ Best case C4 has 30 - 35 percent average reduction.
  - ✓ Case C1 has average 15 percent reduction.
  - ✓ Biased case C0 diverges severely.

35

This chart summarizes observations made in previously exhibited charts. Percent reduction here refers to arithmetically averaged reduction, relative to B0, in the frequency of deterministic CEM MOE results lying outside the four STOCCEM limit bands described in Chart 7. The higher the percent reduction (in divergence) achieved, the better is the case result relative to the SICS objectives.



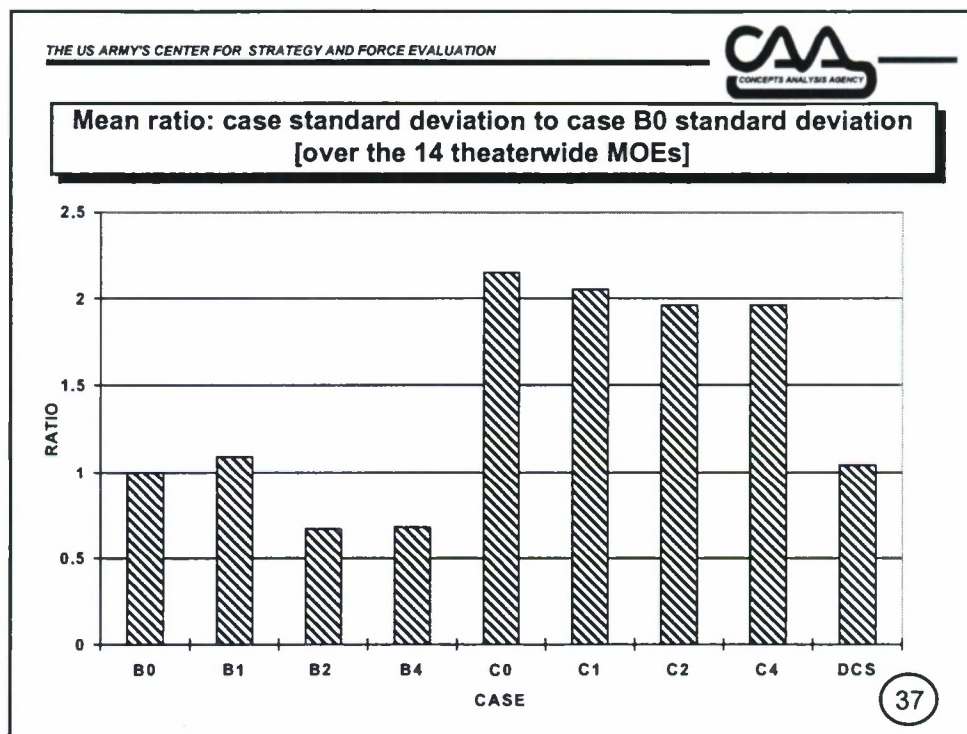
THE US ARMY'S CENTER FOR STRATEGY AND FORCE EVALUATION

CAA  
CONCEPTS ANALYSIS AGENCY

**The following three charts assess STOCCEM variation  
(over replications) in STOCCEM results.**

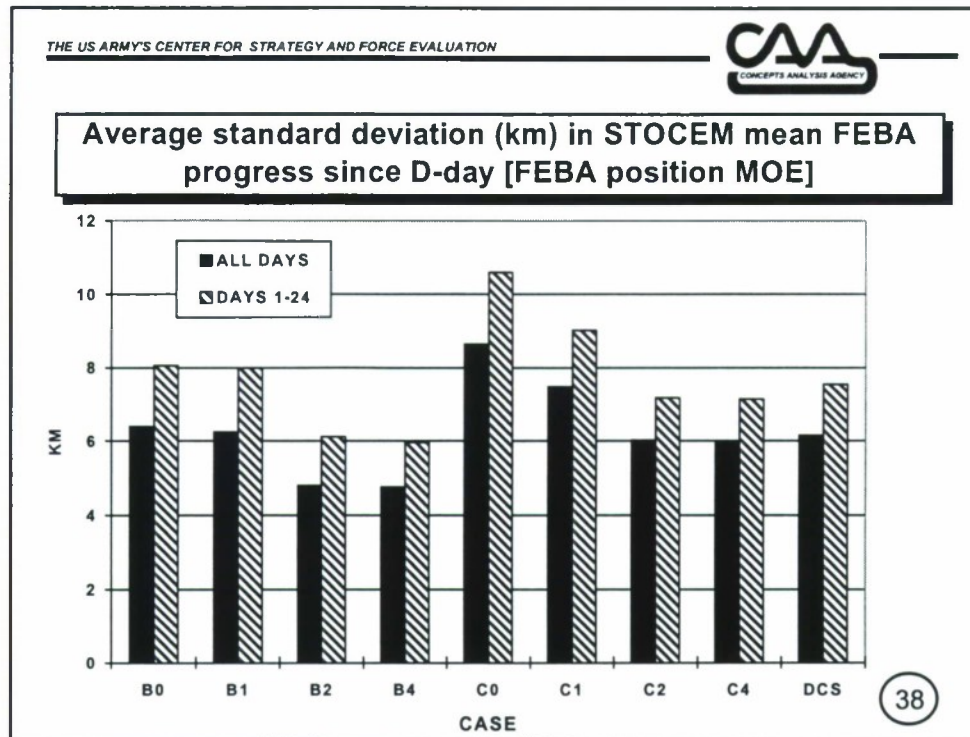
36

Recall, from Chart 5, that one objective of SICS was to seek an increase in variability, over replications, of STOCCEM MOEs, except for those MOEs reflecting FEBA progress.



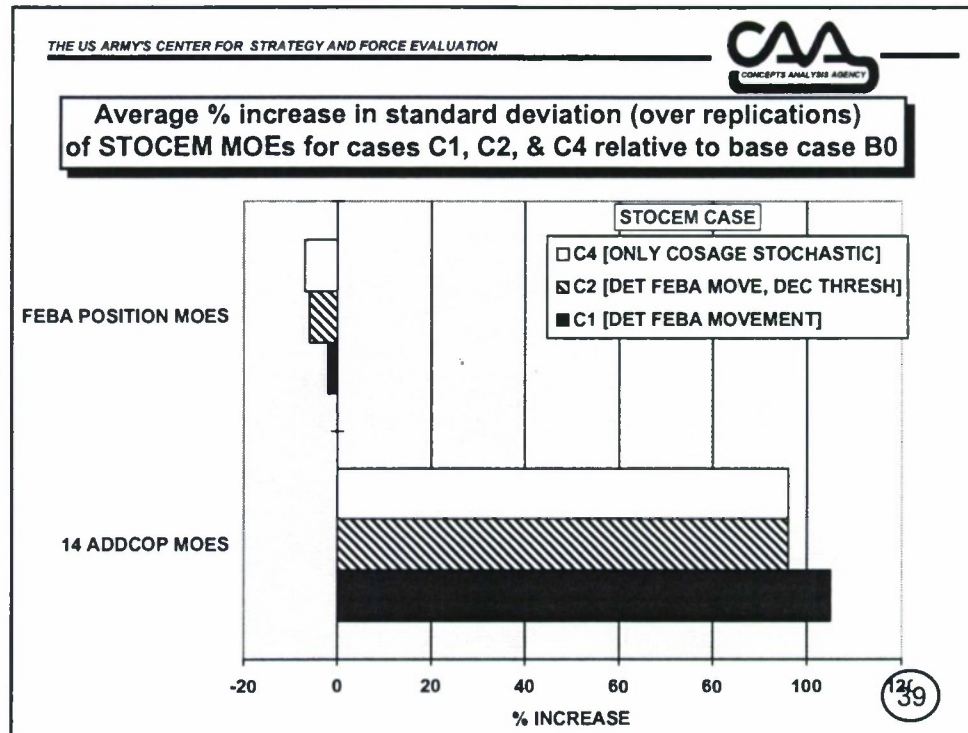
This chart shows, for each STOCCEM case studied, the ratio of the average standard deviation of STOCCEM MOE results for that case to the average standard deviation for the base B0 case. For each of the 14 theaterwide MOE case values in each 4-day theater cycle, the sample standard deviation is computed over the 30 replication results of that MOE. After this standard deviation is computed for all cases, the normalized MOE ratio for each case/theater cycle is computed as ratio of the case standard deviation to the standard deviation for the base case B0 in that theater cycle. For each case, these normalized ratios are then arithmetically averaged over all MOEs and theater cycles to yield the value plotted in this chart.

From the chart, the average STOCCEM variability, over replications, in the 14 theaterwide MOEs, for all of the cases with alternative COSAGE sampling is approximately double that of the base case B0. Cases B2 and B4 show an approximate 33 percent reduction in average variability relative to the base case B0. The B1 and DCS cases have essentially the same average variability as the base case.



This chart shows, for each STOCCEM case studied, the average standard deviation of STOCCEM FEBA position MOE results for that case, in terms of km advanced since D-day (Day 1 of the scenario). For each case, a sample standard deviation was computed for FEBA progress on each of the 21 CEM avenues of advance in each 4-day theater cycle. Each case value plotted on this chart shows the arithmetic average of the standard deviations, computed for that case, over all 21 avenues of advance and all theater cycles in the associated scenario timeframe.

From the chart, differences in average STOCCEM FEBA progress variability, over cases are small. The average standard deviation over the full scenario ranges from around 5 km in cases B2 and B4 to about 8.5 km in case C0.



This chart compares average percent increase in standard deviation (over replications) of STOCES cases results for cases C1, C2, and C4 in Charts 37 and 38, by MOE category, relative to the base case B0. Those cases were designated as preferred based on divergence (deterministic vs stochastic) reduction criteria. Percent increase is arithmetically averaged over all MOEs in a category.

The average standard deviation of FEBA position MOEs for these preferred cases is within 10 percent of that for the base case B0.

For the 14 theaterwide MOEs, the average standard deviation of the preferred case results is essentially double that of the base case.



THE US ARMY'S CENTER FOR STRATEGY AND FORCE EVALUATION

CAA  
CONCEPTS ANALYSIS AGENCY

### Observations on assessment of STOCCEM variability (standard deviation) over replications

- 14 theaterwide MOEs
  - ✓ All cases with alternative COSAGE sampling [C0, C1, C2, and C4] have almost a 100 percent increase in variability in standard deviation over replications (relative to base case).
- FEBA position MOEs
  - ✓ Cases C1, C2, and C4 [using alternative COSAGE sampling] have essentially the same variability in standard deviation over replications as the base case.

40

This chart is self-explanatory and reflects results described in Charts 38 and 39.



## Recommendations

- **Use alternative COSAGE sampling (one posture sample drawn in each replication) to reduce STOCEM divergence from deterministic CEM.**
  - ✓ **Best compromise is to stochastically activate all STOCEM processes except FEBA movement because this option:**
    - has many processes stochastic
    - has only slightly more divergence than activating only the (alternative) COSAGE sampling process.
  - ✓ **Variability in non-FEBA results is also increased.**
- **Do not use stochastic FEBA movement process until it is calibrated to deterministic CEM.**
- **Confirm SICS results with further tests.**

41

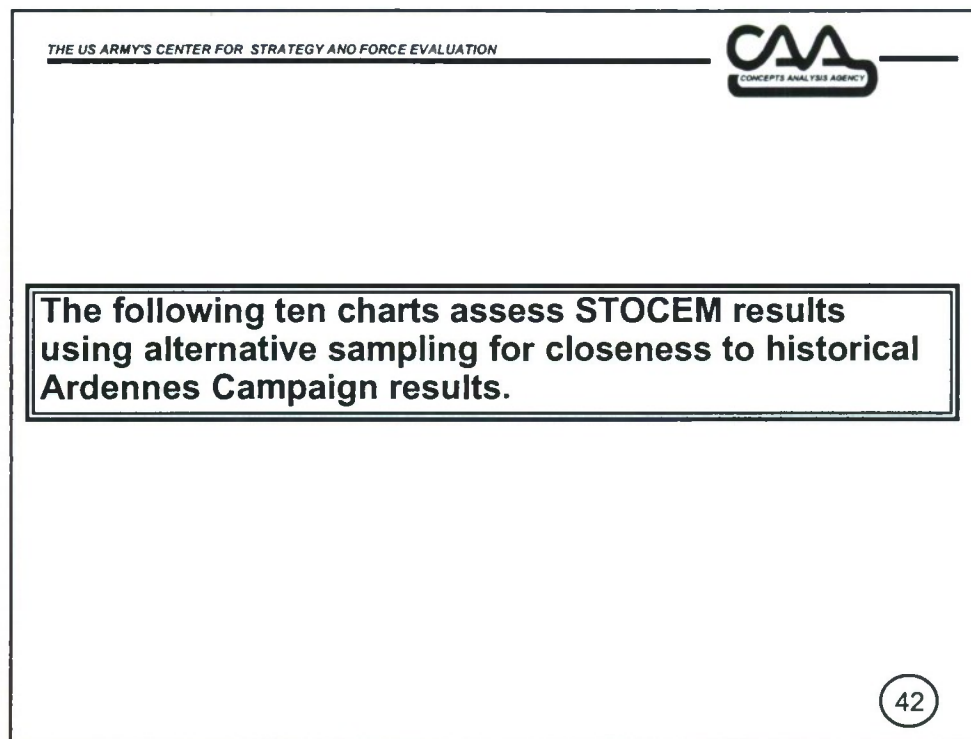
One objective of SICS is to reduce divergence of STOCEM results from deterministic CEM results. Charts 23 and 33 show that, on average, cases C1, C2, and C4, achieve the greatest divergence reductions. C4 showed greater average reduction than C2, and C2 showed greater average reduction than C1. However, the percent reduction was of similar magnitude over these preferred cases. Since C1 applies more stochastic processes than either C2 or C3, it appears to be more desirable than C2 or C4, because it better reflects actual combat, in which virtually all processes are stochastic.

A second objective of SICS is to increase the variability, over replications, of STOCEM results for non-FEBA MOEs, i.e., those MOEs that do not measure FEBA progress. The results in Charts 37 and 38 show that cases C1, C2, and C4 all essentially produce a 100 percent increase in average variability for the 14 theaterwide MOEs, relative to the base case B0, while leaving average variability essentially unchanged for the FEBA progress MOEs.

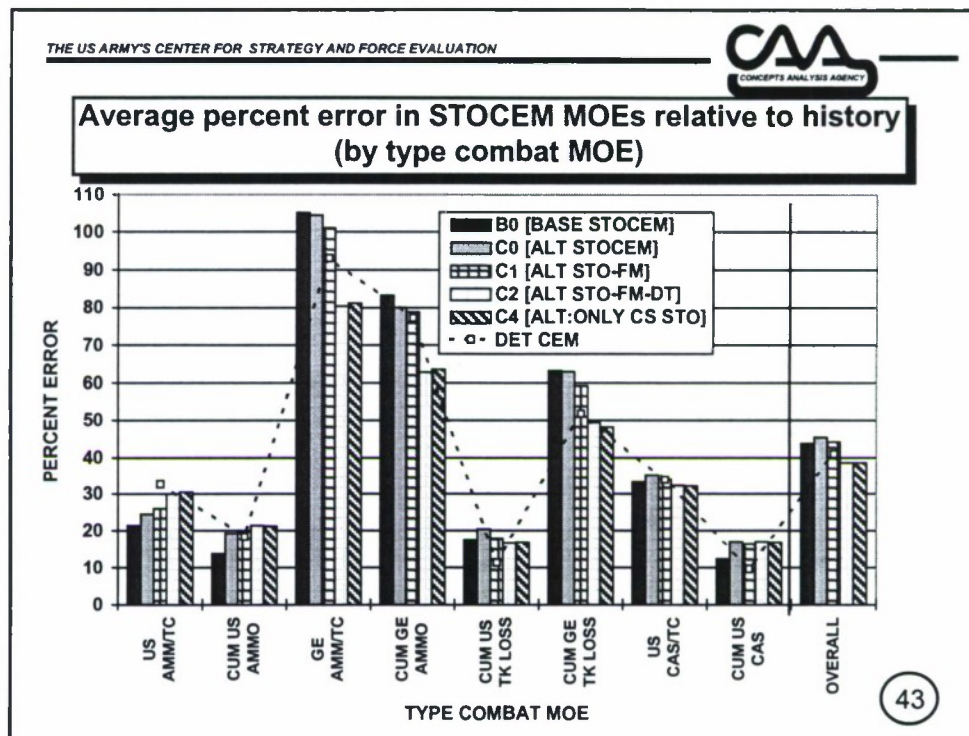
Overall, case C1 appears to be preferred for STOCEM application.

Chart 20 noted that the STOCEM FEBA movement process is biased relative to deterministic CEM because it was not calibrated around deterministic CEM inputs. Because of this bias, STOCEM should not be applied with an active FEBA move rate process until calibration is done.

Since all STOCEM results shown herein are based only on the ARCAS scenario, and reflect a single sample of 30 replications per case, additional experimentation with new scenarios and samples is desirable.



This chart is self-explanatory.

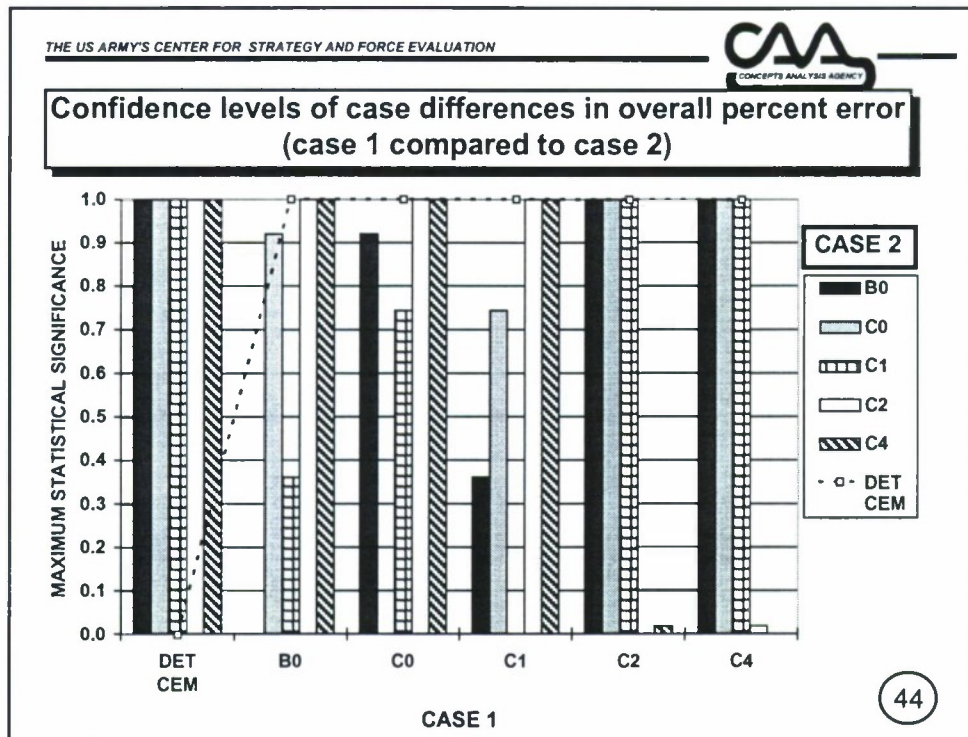


Both historical results and STOCCEM case results for 60 replications were determined at the end of each 4-day theater cycle during the 32-day campaign for the combat MOEs noted in the above chart. Historical results are derived from the Ardennes Campaign Simulation Data Base (ACSDB). The percent error in the STOCCEM mean value relative to history is computed for each MOE value as the absolute value of:  $\left[ \frac{(\text{history} - \text{STOCCEM mean})}{(\text{history})} \right]$  for each STOCCEM replication and theater cycle. The average percent error for an MOE is the arithmetic average of the percent error for the MOE over all eight theater cycle values and all 60 replications. Only eight of the 14 theaterwide MOEs described in Chart 13 are represented in the above chart because historical baseline data have not been compiled for the other six MOEs.

The only cases compared in each case cluster are deterministic CEM (DET CEM), the base case, B0, and those STOCCEM cases in chart 12 which apply alternative COSAGE sampling. The rightmost cluster, labeled OVERALL, plots the arithmetic average of the eight MOE values plotted to the left.

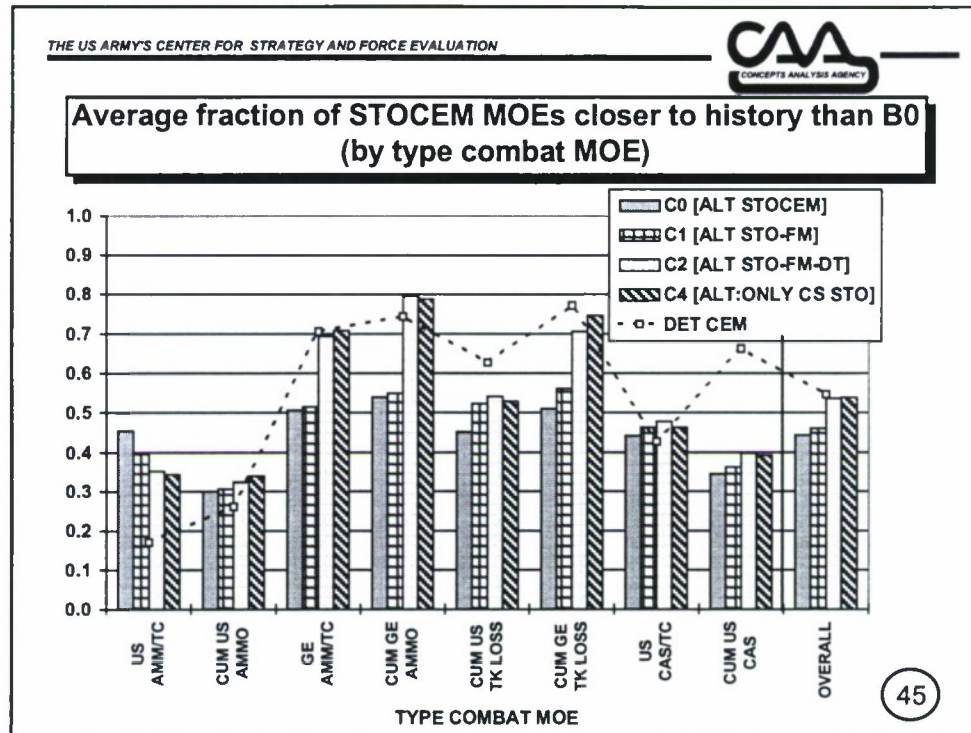
German MOEs produced the largest average discrepancies (percent error) between history and STOCCEM. Overall, cases C2, and C4 qualify as best cases, having the lowest average discrepancies from history (OVERALL percent error of 39 percent each). Case C0 is the worst case, having the largest average discrepancy (45 percent). All other cases have average OVERALL percent error between 41 percent and 44 percent. Statistical confidence level of OVERALL differences between STOCCEM cases is discussed in the next chart.





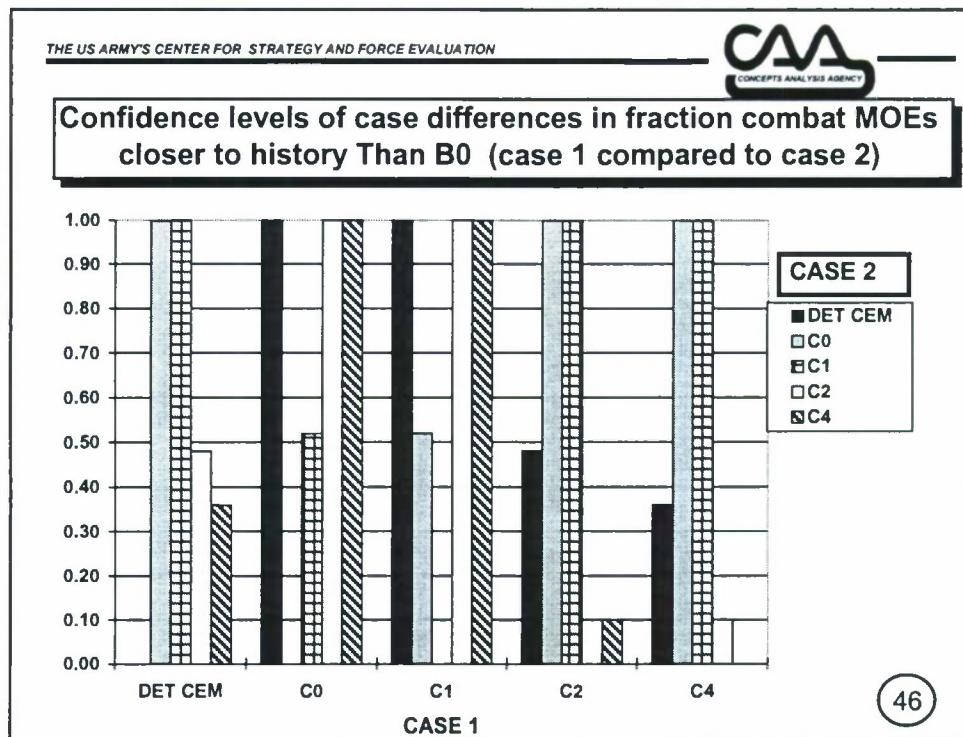
This chart treats the statistical confidence level of differences between paired OVERALL means shown in the previous chart. Each case 1 OVERALL mean (MOE percent error relative to history) was subtracted from each case 2 OVERALL mean. This difference in case means was then tested for the statistical confidence associated with accepting the hypothesis that it is different from zero, i.e., that the means represent different statistical populations. For a sample of size 60 (replications), the sampling distribution of differences in means can be treated as normally distributed. The resulting statistical confidence levels associated with (accepting) the difference between case means are plotted in the above chart. The closer a confidence level is to 1.00, the higher the level of assurance that the associated difference is statistically significant. For example, the leftmost bar of the "CASE 1 = B0" cluster indicates that the difference between the case C0 OVERALL mean and the case B0 OVERALL mean, as shown in the OVERALL cluster of Chart 43, is statistically significant at the .92 confidence level. This is the same as the ordinate of the B0 (leftmost) bar in the cluster for case 1 equal to C0.

This chart supports the best case/worst case designations made in Chart 43. Confidence level test results indicate that, relative to OVERALL means shown in chart 43, the best cases, C2 and C4, are not significantly different from each other, and these cases both differ from cases B0, C1, C0, and DET CEM, with a confidence level above .99. The worst case, C0, differs from case C1 with .74 confidence level, and from all other cases with at least .92 confidence.



For the specified SICS STOCCEM cases, this chart shows, for each STOCCEM case, the fraction of MOE values which are closer to history (based on the ACSDB) than corresponding MOE values from the base case, B0. Closeness to history for an MOE type is in terms of the absolute difference between history and a corresponding STOCCEM result, and is assessed, and averaged, over individual theater cycle results of 60 STOCCEM replications. The MOE types are the same as in Chart 43. The cases compared are deterministic CEM, and those STOCCEM cases in Chart 12 which apply alternative COSAGE sampling. The rightmost cluster, labeled OVERALL, plots the arithmetic average of the eight MOE values plotted to the left.

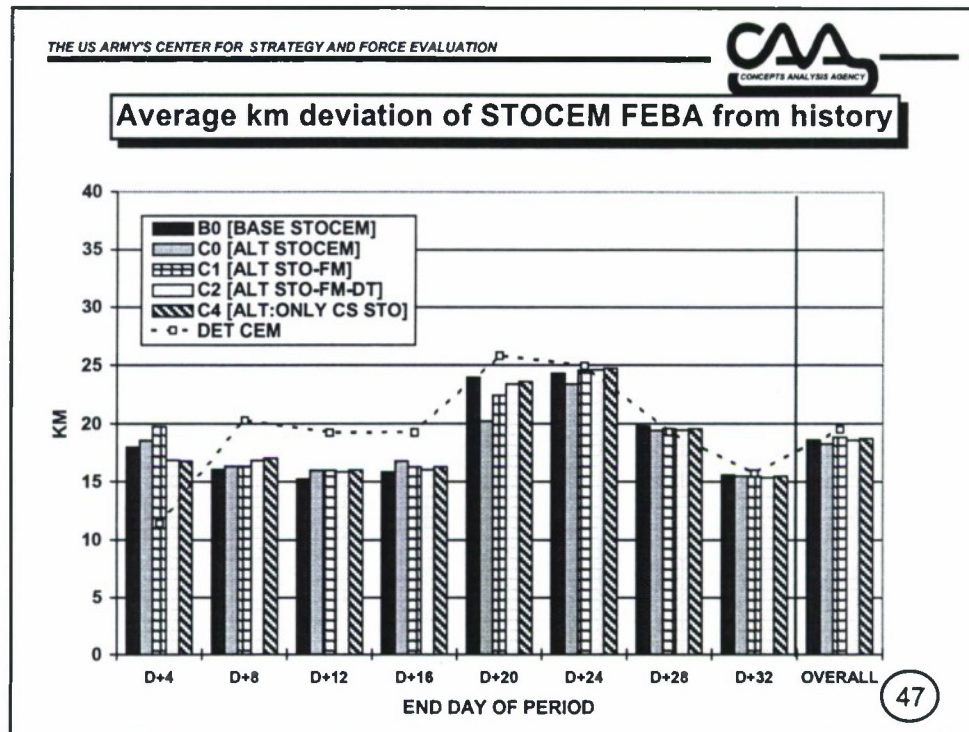
Overall, deterministic CEM qualifies as the best case, having the largest OVERALL fraction (55 percent) of MOE values closer to history than base case B0. Case C0 is the worst case, having only 44 percent closer to history than B0. Only cases C2, C4, and deterministic CEM, with overall fractions exceeding .50, are closer to history than the base case B0. Statistical confidence levels associated with OVERALL differences between STOCCEM cases are discussed in the next chart.



This chart treats the statistical confidence level associated with differences between paired OVERALL means (for fraction closer to history than B0) shown in the previous chart. Exactly analogous to Chart 44, each difference (case 1 OVERALL mean vs case 2 OVERALL mean) was tested for the statistical confidence level associated with accepting the hypothesis that it is different from zero, i.e., that the means represent different statistical populations. The sampling distribution of differences in means is again treated as normally distributed. The confidence levels associated with (accepting) the differences between paired case OVERALL means are plotted in the above chart.

This chart supports the best case/worst case designations made in Chart 45. Confidence level test results indicate that, relative to OVERALL means depicted in Chart 45, the best case, deterministic CEM, is not significantly different from cases C2 and C4, and these cases both differ from cases C0 and C1 with a confidence level above .99. The worst case, C0, is not statistically distinguishable from case C1, since the difference has only a .52 confidence level. However, both case C0 and C1 differ from the other three cases with a confidence level exceeding .99.

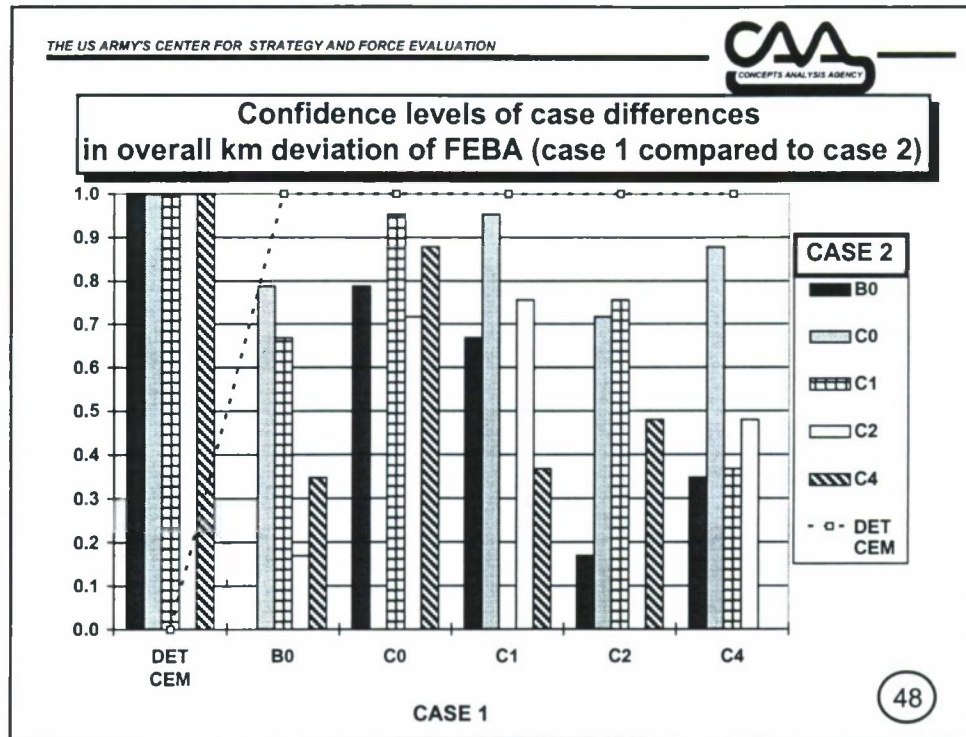




For selected SICS cases, both historical results and STOCCEM FEBA position MOE results were determined for each of 21 avenues of advance in the STOCCEM theater at the end of each 4-day theater cycle. The absolute difference between history and a STOCCEM case result, is in terms of km advanced since D-day. The average km deviation from history for each 4-day period was computed as the arithmetic average, over 60 STOCCEM replications, of the absolute differences, in each 4-day period, over all 21 avenues of advance. The only cases compared in each cluster are deterministic CEM, the base case, B0, and those STOCCEM cases in chart 12 which apply alternative COSAGE sampling. The rightmost cluster, labeled OVERALL plots the arithmetic average of the eight MOE values plotted to the left.

Case C0 qualifies as the the best case, having the smallest average OVERALL deviation (from history) of 18.3 km. Deterministic CEM is the worst case, having the largest discrepancy between history and STOCCEM, an OVERALL deviation of 19.5 km. Statistical confidence levels associated with OVERALL differences between STOCCEM cases are discussed in the next chart.

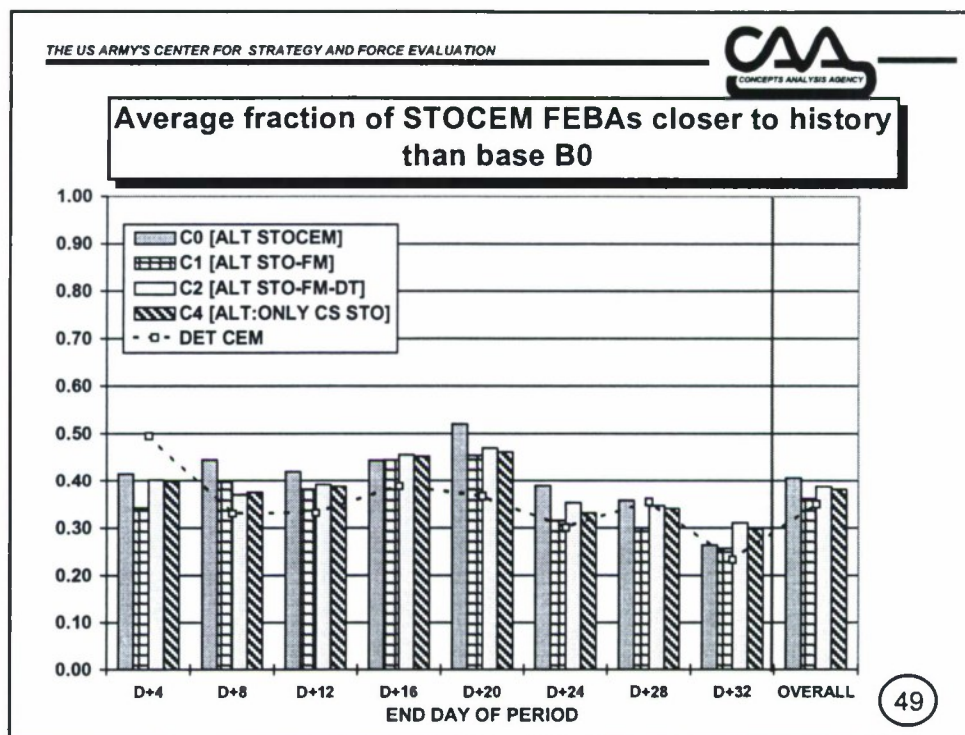




This chart treats the statistical confidence level of differences between paired OVERALL means (for km deviation from history) shown in the previous chart. Each difference (case 1 OVERALL mean vs case 2 OVERALL mean) was assessed for the statistical confidence level associated with accepting the hypothesis that it is different from zero, i.e., that the means represent different statistical populations. The sampling distribution of differences in means is again treated as normally distributed. The resulting confidence level associated with (accepting) the difference between paired case OVERALL means is plotted in the above chart.

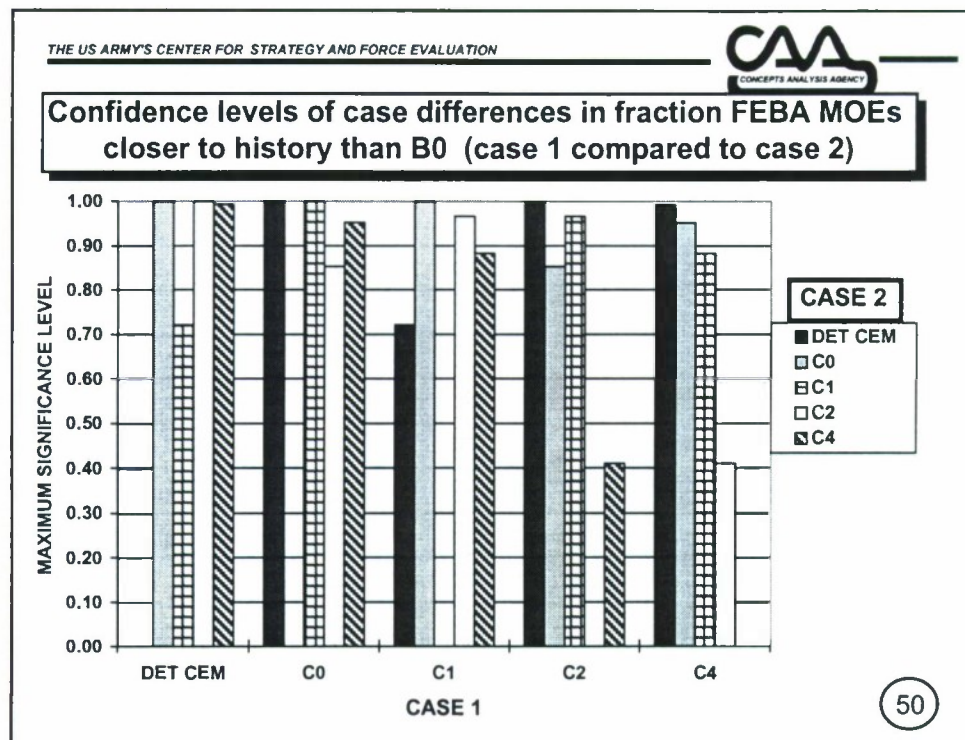
These results indicate that, relative to OVERALL means depicted in chart 47, the assessed best case, C0, differs from case C2 with a .72 confidence level, from case B0 with a .79 confidence level, and from cases C1, C4, and deterministic CEM with at least a .88 confidence level. This provides qualified support to the best case designation assigned to C0 in Chart 47. The assessed worst case, deterministic CEM, differs from all other cases with at least a .99 confidence level. This strongly supports the worst case assessment assigned to deterministic CEM in Chart 47.

Case B0 and the best case, C0, are statistically similar (differing only at the .79 confidence level). B0 and C0 are also the only cases using the stochastic FEBA move process. These results therefore suggest that the STOCCEM stochastic FEBA move process may be a better one, relative to history, than is the FEBA movement process used in deterministic CEM.



Exactly 168 FEBA position MOE values (for each of 21 avenues of advance and of eight 4-day periods) were computed for each FEBA position MOE for each STOCCEM case and for the history baseline (from the ACSDB). For the specified STOCCEM cases, this chart shows the fraction of each case's FEBA position MOE values which are closer to history than the corresponding values for the base case B0. Closeness to history is in terms of the absolute difference between history and a corresponding STOCCEM result, and is assessed, and averaged, over individual theater cycle results of 60 STOCCEM replications. The cases compared are deterministic CEM, and those STOCCEM cases in chart 12 which apply alternative COSAGE sampling. The rightmost cluster, labeled OVERALL plots the arithmetic average of the eight MOE values plotted to the left.

Overall, case C0 is the best case, having the largest OVERALL fraction (.41) closer to history than base case B0. Deterministic CEM (.35) is the is worst case. All other cases have average OVERALL percent error between .36 and .39. All depicted fractions are considerably less than .50 because approximately 14 percent of STOCCEM MOES are equal (to zero) over all cases. This occurs because the STOCCEM FEBA retreat along each avenue of advance is constrained to stop at the D-day FEBA positions (0 km advance) in all cases. Statistical confidence levels of OVERALL differences between STOCCEM case means are discussed in the next chart.



This chart treats the statistical confidence level of differences between paired OVERALL means (for fraction closer to history than B0) shown in the previous chart. Each difference (case 1 OVERALL mean vs case 2 OVERALL mean) was tested for the statistical confidence level associated with accepting the hypothesis that it is different from zero, i.e., that the means represent different statistical populations. The sampling distribution of differences in means is again treated as normally distributed. The resulting confidence levels associated with (accepting) the difference between paired case OVERALL means are plotted in the above chart.

This chart supports the best case/worst case designations made in Chart 49. Confidence level test results indicate that, relative to OVERALL means depicted in Chart 49, the best case, C0, statistically differs from case C2 with a .85 confidence level, and from all other cases with at least a .95 confidence level. The worst case, deterministic CEM, differs from case C1 with only a .72 confidence level, but differs from all other cases with a confidence level exceeding .99. Cases C2 and C4 appear to be very similar statistically (associated confidence level of difference = .41). Some uncertainty might be resolved by additional sampling.



THE US ARMY'S CENTER FOR STRATEGY AND FORCE EVALUATION



### Summary observations on closeness to history over the eight combat MOES


- **Best case results (closest to history):**
  - ✓ Cases C2 and C4 have lowest average percent deviation from history.
    - Difference from other cases is significant (>.99 confidence)
  - ✓ Deterministic CEM has the largest fraction of MOE values which are closer to history than base B0.
    - Difference from C2 and C4 is not statistically significant.
    - Difference from C0 and C1 is significant (>.99 confidence).
- **Worst case results (furthest from history):**
  - ✓ Case C0 has largest average percent deviation from history.
    - Difference from B0 and C1 has modest (.74-.92) confidence.
    - Difference from C4, C2, and DET CEM is significant (>.99) .
  - ✓ Case C0 has lowest fraction MOEs closer to history than B0.

51

This chart summarizes results already described in Charts 43 through 46. Using a “closeness to history” criterion with the OVERALL combat MOE results, STOCCEM cases C2 and C4 were preferred since they had the lowest overall percent deviation from history. Case C2 and C4 were also jointly preferred, along with deterministic CEM, under the criterion of having the largest fraction of combat MOE values which are closer to history than those for the base case, B0.

Case C0 was the least preferred case under both criteria. C0 was statistically similar to case C1, but differed with high confidence from other cases.



THE US ARMY'S CENTER FOR STRATEGY AND FORCE EVALUATION 

### Summary observations on closeness to history over the FEBA position MOEs

- **Best case results (closest to history):**
  - ✓ Case C0 has lowest average divergence from history.
    - Difference from B0 and C2 weakly significant (<.79 conf).
    - Difference from cases C1, C4, and deterministic CEM is significant (.88 - .99 confidence level).
  - ✓ Case C0 has largest fraction MOEs closer to history than B0.
- **Worst case results (furthest from history):**
  - ✓ Deterministic CEM case has largest divergence from history.
    - Statistically different from all other cases (>.99 confidence).
  - ✓ Deterministic CEM has lowest fraction MOEs closer to history than base B0.
    - Difference from C1 is only weakly significant (.72 conf).
    - Statistically different from other cases (>.99 confidence).

52

This chart summarizes results already described in Charts 47 through 50. Using a “closeness to history” criterion with the OVERALL FEBA position MOEs, STOCCEM case C0 was most preferred, having the lowest overall percent deviation from history, but this case was not strongly differentiated from cases B0 and C2. C0 was also the best case under the criterion of having the largest fraction of combat MOE values closer to history than the base case, B0. Under this criterion, C0 differs from case C2 with .85 confidence and from all other cases with at least .95 confidence.

Deterministic CEM was the least preferred case under both criteria. This nonpreference was strongly differentiated statistically (>.99 confidence) in almost all cases. Deterministic CEM differed weakly (.72 confidence) from case C1 in the overall fraction of combat MOE values closer to history than the base case, B0.

THE US ARMY'S CENTER FOR STRATEGY AND FORCE EVALUATION



## Consolidated observations and recommendations

- Initial (Chart 41) recommendation was to use case C1:
  - ✓ Use alternative COSAGE sampling (1 posture sample drawn in each replication) to reduce STOCCEM divergence from deterministic CEM.
  - ✓ Stochastically activate all STOCCEM processes except FEBA movement.
  - ✓ Cases C1, C2, and C4 all have similar preference relative to divergence from deterministic CEM and output variability.
- History comparison prefers use of alternative case C2.
  - ✓ Stochastically activate all STOCCEM processes except FEBA movement and decision thresholds.
  - ✓ C2 has lowest average percent error over eight combat MOEs.
- History FEBA comparison suggests calibrating deterministic CEM FEBA move process to stochastic STOCCEM process.

53

This chart summarizes results shown in Charts 51 and 52 and uses them to qualify the recommendations in Chart 41. Based on the objectives of reducing STOCCEM divergence from deterministic CEM and of increasing variability in non-FEBA results, cases C1, C2, and C4 are almost equally preferred. Chart 41 recommends use of case C1 chiefly because it activates the largest number of stochastic processes. The historical comparison suggests, with high statistical confidence, that case C2 is most preferred because it produces the smallest percent deviation from history over the eight combat MOEs in the scenario examined. Case C2 uses alternative COSAGE sampling, but does not use a stochastic FEBA movement process or stochastic decision threshold processes.

Additionally, the comparison with historical FEBA progress indicates, with only modest statistical confidence, that STOCCEM FEBA results with the stochastic FEBA movement process are closer to history than results using the deterministic CEM FEBA movement process. Since the STOCCEM FEBA movement process was not calibrated around the deterministic CEM process, these results suggest that deterministic CEM should perhaps be “backward-calibrated” to reflect the stochastic FEBA movement process of STOCCEM.

Further investigation may be warranted in view of the limited nature of the scenario and samples examined. Testing and assessment with additional sample replications may increase confidence differences that were only weakly significant in these results.

**APPENDIX A**  
**RAA CONTRIBUTORS**

**RAA TEAM**

**a. RAA Director**

Mr. Walter J. Bauman, Tactical Analysis Division

**b. Team Member**

Mr. William Allison, Operations Support Division

**c. Contributor**

Dr. Ralph Johnson, Operations Support Division

## APPENDIX B

## REQUEST FOR ANALYTICAL SUPPORT

P A R T  1	<b>REQUEST FOR ANALYTICAL SUPPORT</b>			
	1. Performing Directorate/ Division: TA		2. Account Number: 97093	
	3. Type Effort (Enter one):  Mode (Contract=C) <input type="checkbox"/> <input checked="" type="checkbox"/> R		4. Tasking (Enter one):  <input type="checkbox"/> I <input checked="" type="checkbox"/> F - Formal Directive I - Informal V - Verbal	
	S - Study Q - QRA P - Project R - RAA M - MMS			
	5. Title: STOCER Investigation of COSAGE Sampling			
	6. Acronym: SICS		7. Date Request Received: 02/28/97	
	8. Date Due: 08/30/97		9. Requester/Sponsor (i.e., DCSOPS): CAA	
	10. Sponsor Division (i.e., SSW, N/A) TA/28/97		11. Impact on Other Studies, QRA, Projects, RAA: Concurrent with and supporting KOSAVE study	
	12. Product Required: Briefing			
	13. Estimated Resources Required:		a. Estimated PSM: 2.5	
b. Estimated Funds:		c. Models Req'd: STOCER		
d. Other:		14. Objective(s)/Abstract: Combat samples generated by COSAGE were input to STOCER in the Ardennes Campaign Simulation (ARCAS) study. Planning for use of combat samples in the Kursk Operation Simulation and Validation (KOSAVE) study will examine several methods for using ARCAS COSAGE samples in STOCER. In support of this planning, SICS will execute STOCER with several different stochastic sampling techniques applied to the COSAGE samples used in ARCAS, and will compare average STOCER outcome results and ranges of outcome variation. Results will guide the use of COSAGE samples in KOSAVE and will be of interest to all users of STOCER.		
15. Study Director/POC:		Last Name: Bauman		
First: Walter		Date: 02/28/97		
Signature: <i>Walter A Bauman</i>		Phone#: 295-5261		
GO TO BLOCK 20 If this is A STUDY. See Tab C of the Study Directors' Guide for preparation of a Formal Study Directive.				
P A R T  2	16. Background/Statement of Problem*: The range of variation in STOCER outcomes in ARCAS was often small. There is a need to investigate whether variability in COSAGE samples is insufficiently represented in the ARCAS methodology in order to plan for effective use of COSAGE samples in the KOSAVE study.			
	17. Scope of Work*: Formulation of sampling methods to be studied. Construction of STOCER inputs to sample ARCAS COSAGE samples using these methods. Formulation of MOE's studied. Modification of STOCER output processors to process results. Execution of 30 replications of STOCER for each method. Comparison, analysis, and documentation of results from each method.			
	18. Issues for Analysis*: How does the average STOCER result vary using each COSAGE sampling method? How does the range of STOCER outcome results vary using each sampling method? What sampling method is preferred?			
	19. Milestones/Plan of Action*: Define problem and methods - 15 March. Develop STOCER inputs and modify postprocessors - 15 April. Complete STOCER executions 15 May. Complete analysis - 1 July. Brief ARB 30 July.			
	20. Division Chief Concurrence:		Date: 28/2/97	
	21. Sponsor (COL/DA Div Chief) Concurrence:		Date:	
22. Sponsor Comments*:				



## **APPENDIX C**

### **REFERENCES**

1. Stochastic Concepts Evaluation Model (STOCEM), CAA-TP-91-6, US Army Concepts Analysis Agency, August 1991
2. Stochastic Concepts Evaluation Model - Phase II (STOCEM-PHASE II), CAA-TP-92-2, US Army Concepts Analysis Agency, June 1992
3. Stochastic Concepts Evaluation Model - Phase II I(STOCEM-PHASE III), CAA-TP-93-2, US Army Concepts Analysis Agency, June 1993
4. Ardennes Campaign Simulation (ARCAS), Study Report CAA SR-95-8, US Army Concepts Analysis Agency, December 1995, ADA 307 014
5. Concepts Evaluation Model IX (CEM IX), Volume II - User's Handbook, CAA-D-85-1, US Army Concepts Analysis Agency, August 1985 (revised December 1995)

## APPENDIX D

## HISTORY VS STOCCEM MEAN COMBAT MOEs IN EACH 4-DAY PERIOD

**OVERVIEW.** Figures D-1 through D-8 show historical results and mean STOCCEM results averaged over two samples of 30 replications at the end of each 4-day theater cycle for eight combat MOEs. Only eight of the 14 theaterwide MOEs described in Chart 13 were suitable for comparison because historical baseline data had not been compiled for the other six MOEs. The cases depicted are deterministic CEM, the base STOCCEM case, B0, and STOCCEM cases C0, C1, C2, and C4, which apply alternative COSAGE sampling.

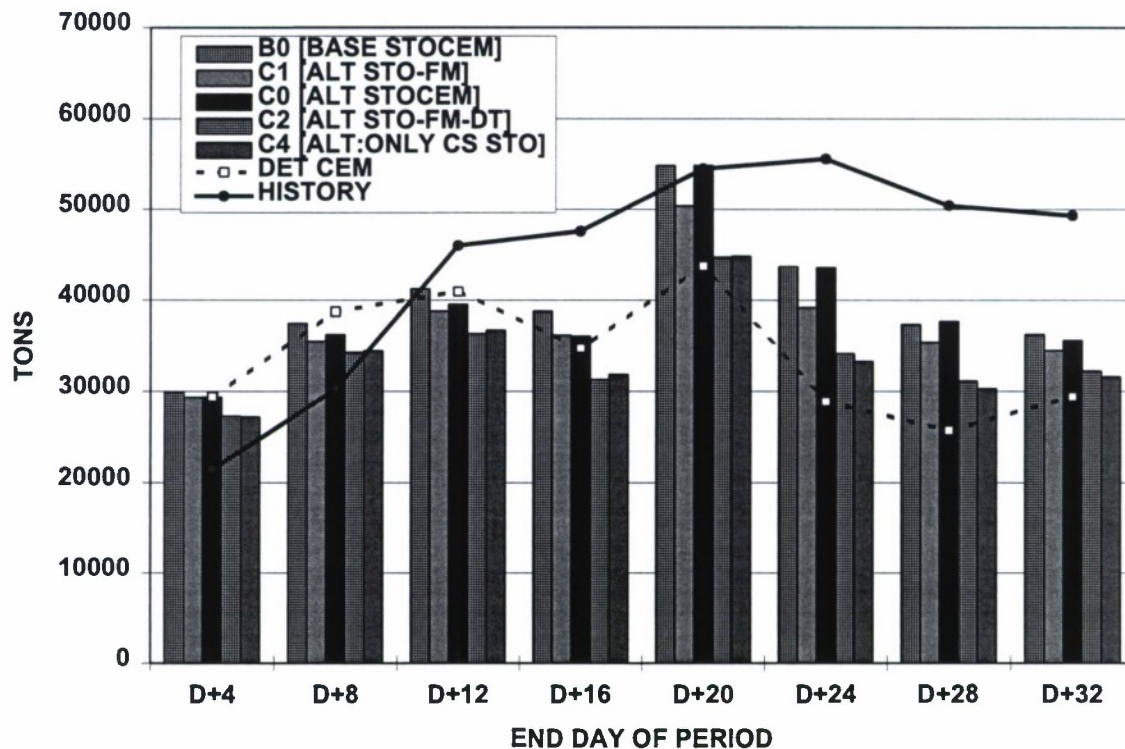


Figure D-1. Average US/UK Ammunition Consumed in Each Period

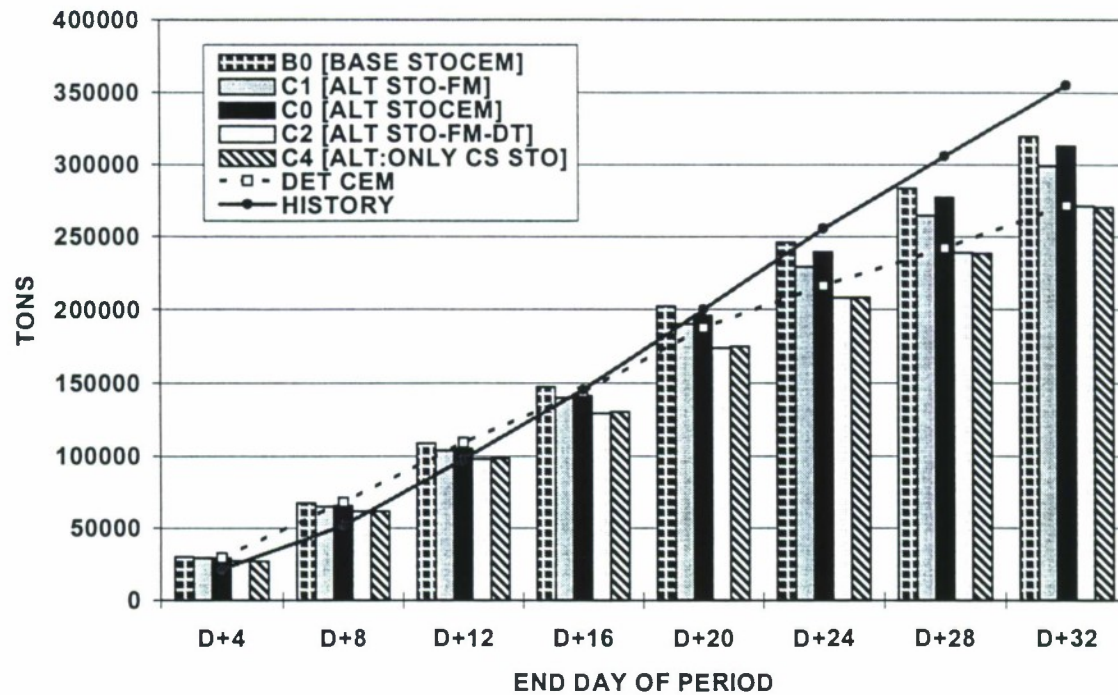


Figure D-2. Average Cumulative US/UK Ammunition Consumed

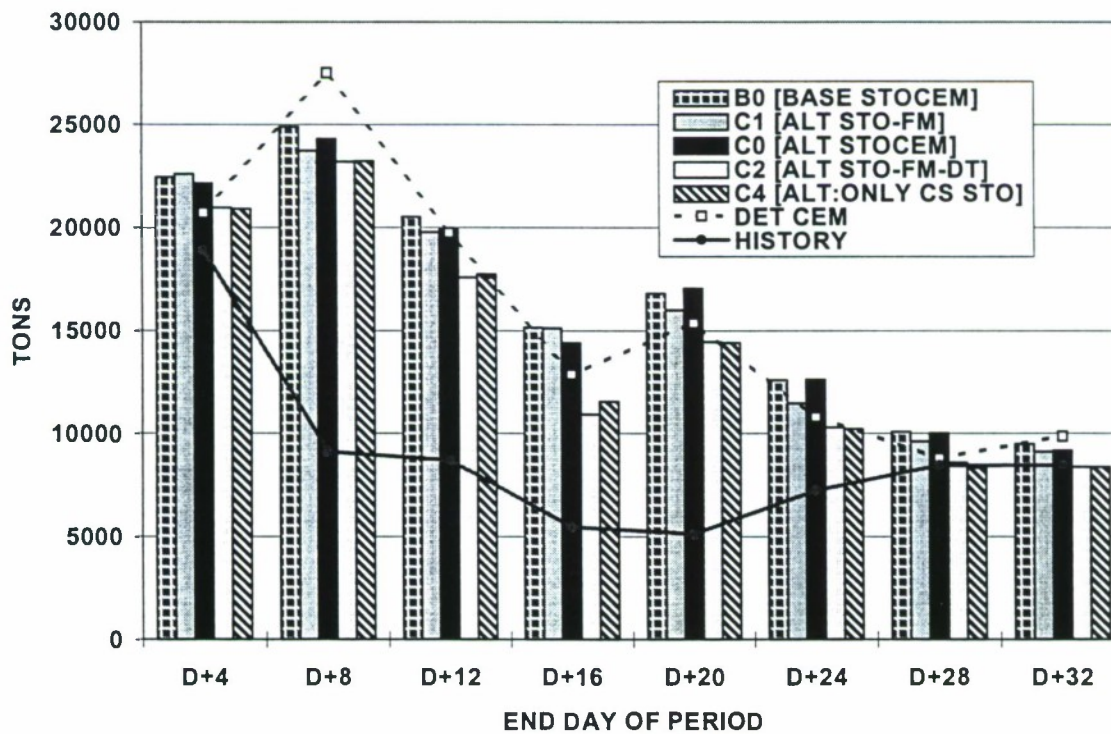


Figure D-3. Average German Ammunition Consumed in Each Period



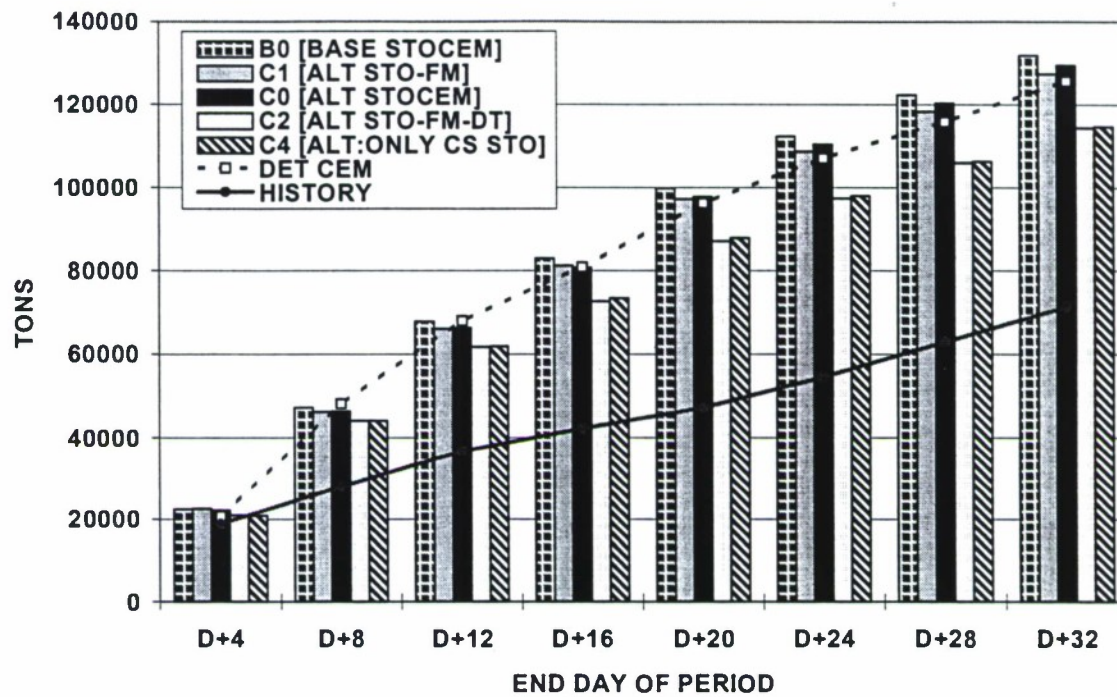


Figure D-4. Average Cumulative German Ammunition Consumed

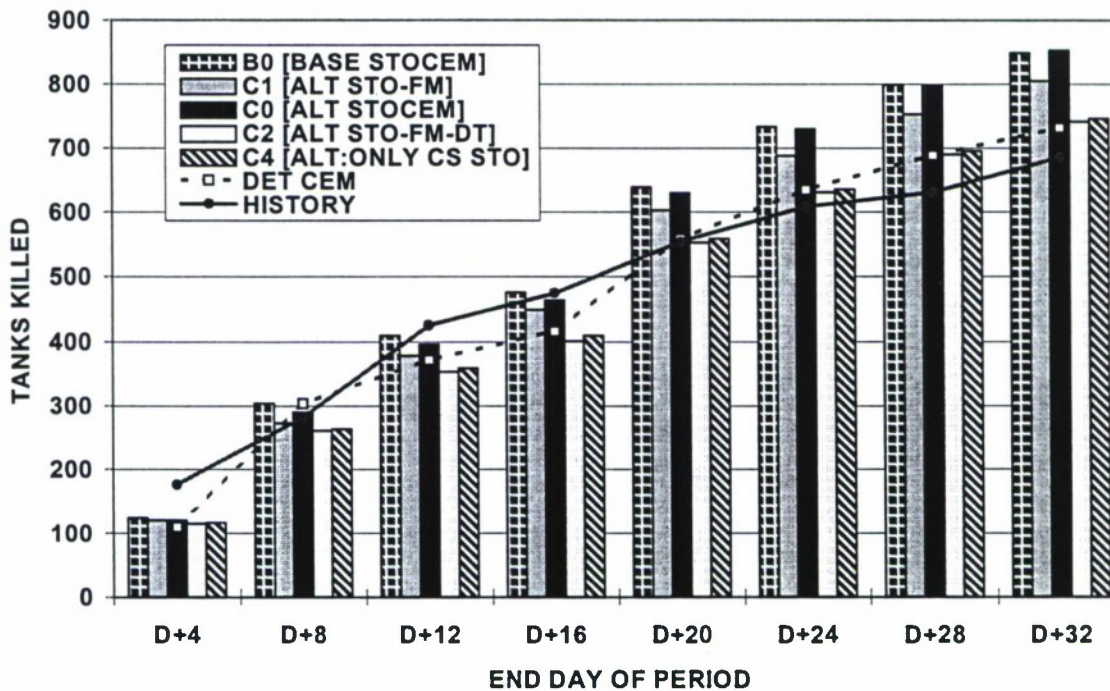


Figure D-5. Average Cumulative US/UK Permanent Tank Losses



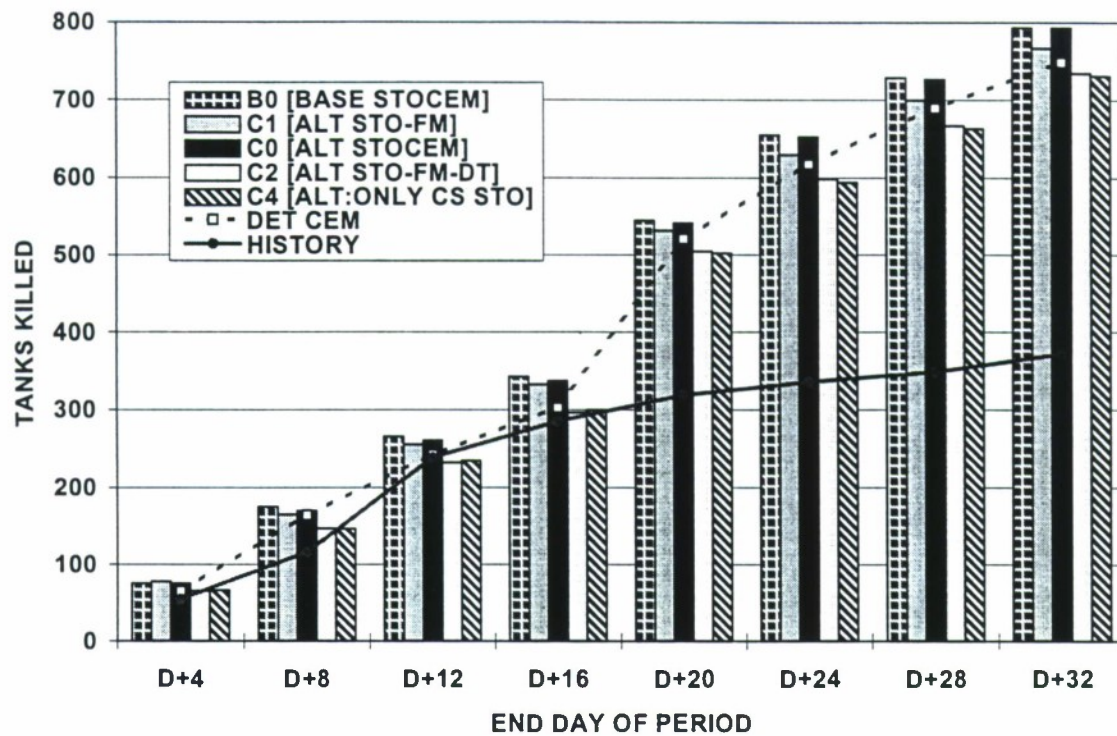


Figure D-6. Average Cumulative German Permanent Tank Losses

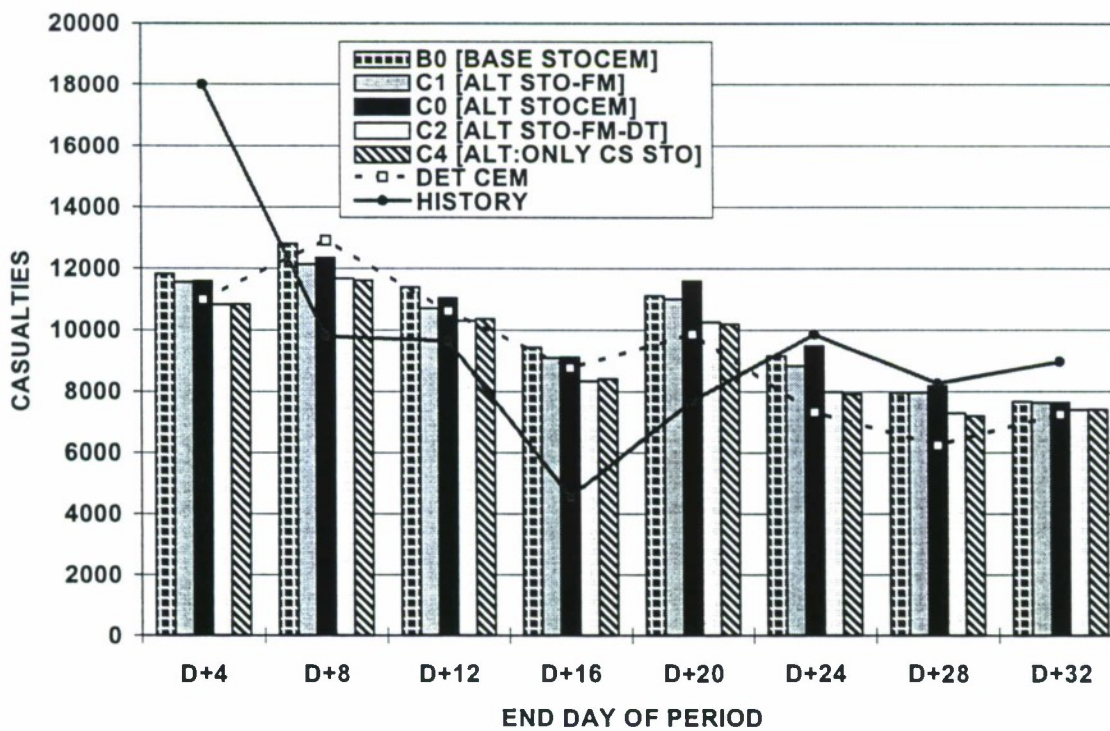


Figure D-7. Average US/UK Permanent Personnel Casualties in Each Period

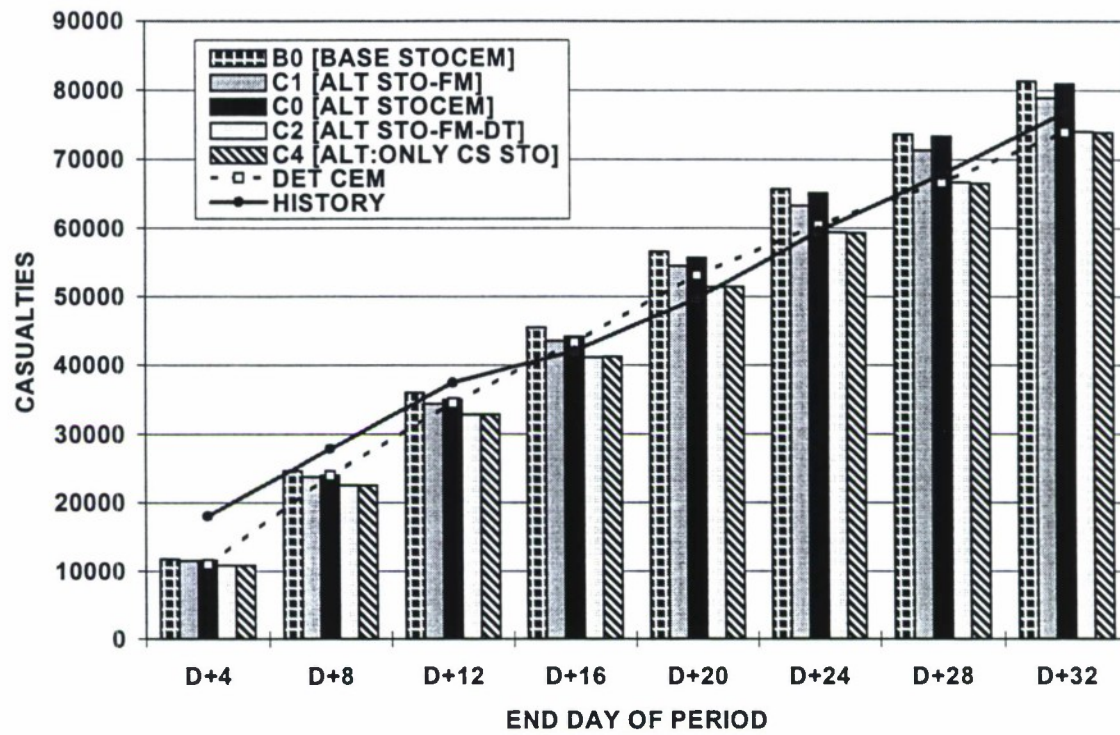


Figure D-8. Cumulative US/UK Permanent Personnel Casualties in Each Period

## APPENDIX E

## HISTORY VS STOCCEM MEAN FEBA POSITION MOEs IN EACH 4-DAY PERIOD

**OVERVIEW.** Figures E-1 through E-8 show historical results and mean STOCCEM FEBA progress results, averaged over two samples of 30 replications, along each STOCCEM avenue of advance at the end of each 4-day theater cycle. The 21 avenues of advance are the same as were used in ARCAS. For modeling purposes, the STOCCEM theater area is overlaid with a system of 21 movement corridors, denoted herein as CEM “avenues of advance,” corresponding to the initially planned flow of forces in STOCCEM during the campaign. These avenues are serially indexed in a north-to-south geographic ordering as avenue 10 through avenue 30. The magnitude of the cumulative FEBA progress since D-day (Day 1 of the scenario) is plotted for each avenue of advance. The D-day position is at the 0 ordinate, and a negative “km from D-day FEBA” corresponds to a German advance. This linearized representation emulates a quasi-geography for the battle with relative positions along the (north-south ordered) STOCCEM avenues of advance represented as parallel straight lines. The orientation is from an aerial perspective facing east from above US/UK lines. The cases depicted are deterministic CEM, the base STOCCEM case, B0, and STOCCEM cases C0, C1, C2, and C4, which apply alternative COSAGE sampling.

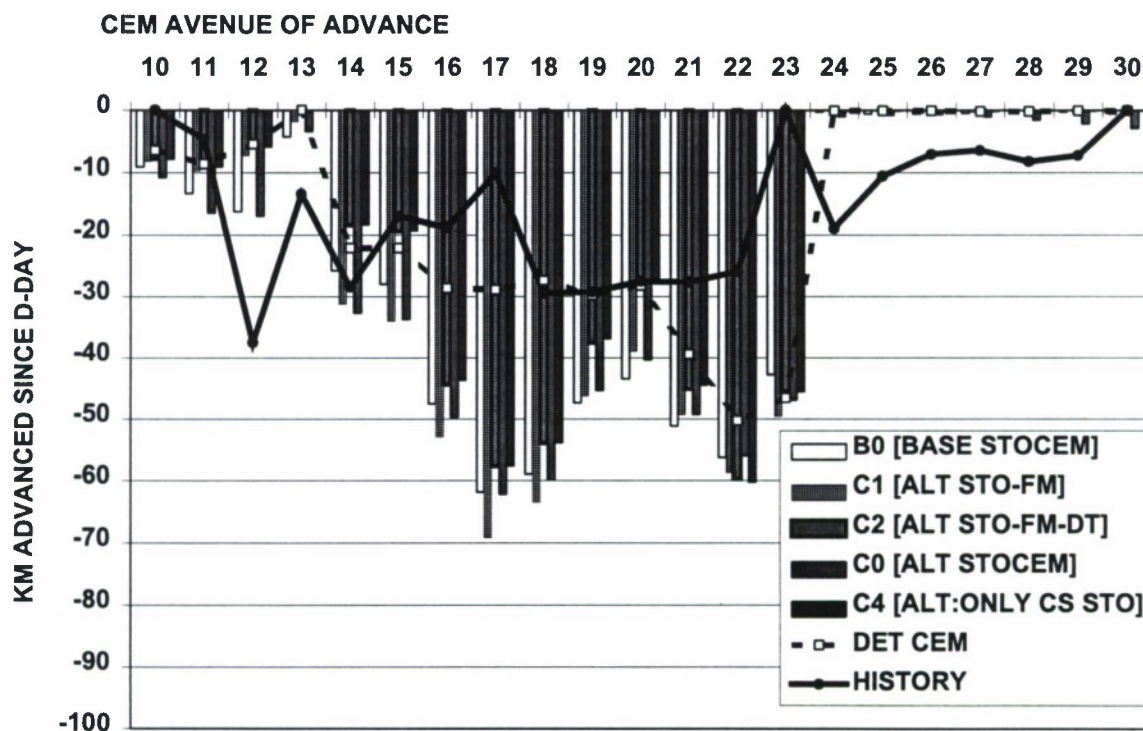


Figure E-1. FEBA Positions on D+4



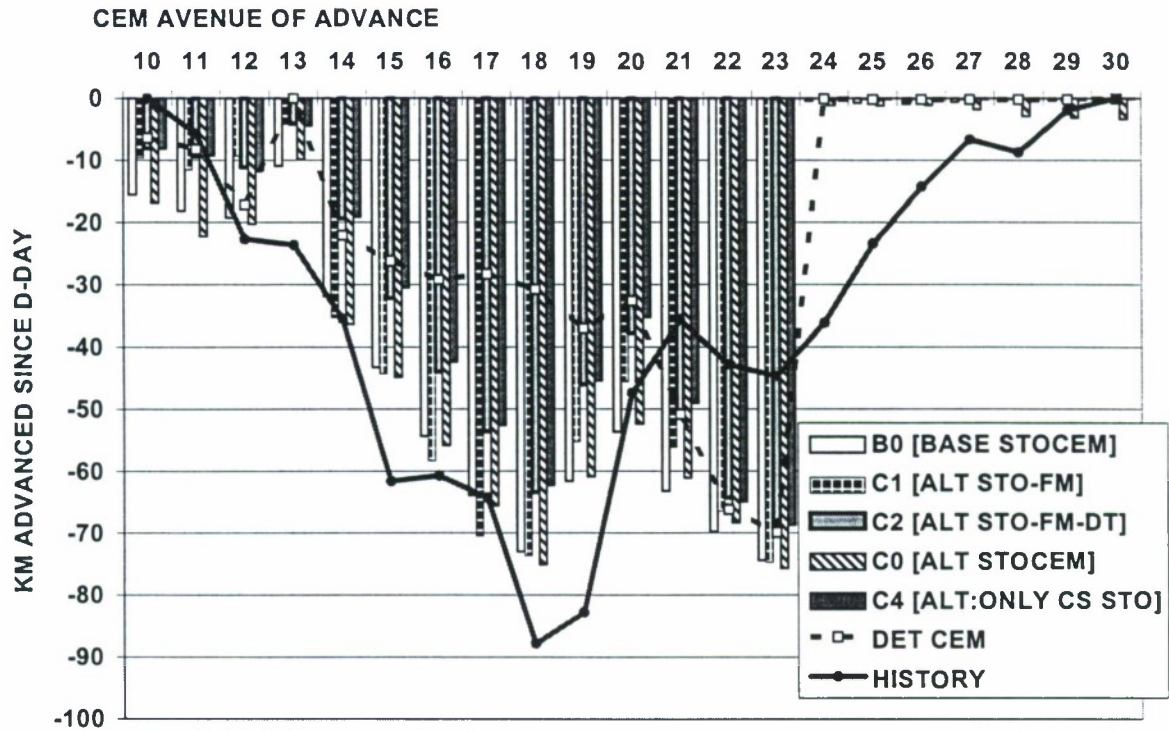


Figure E-2. FEBA Positions on D+8

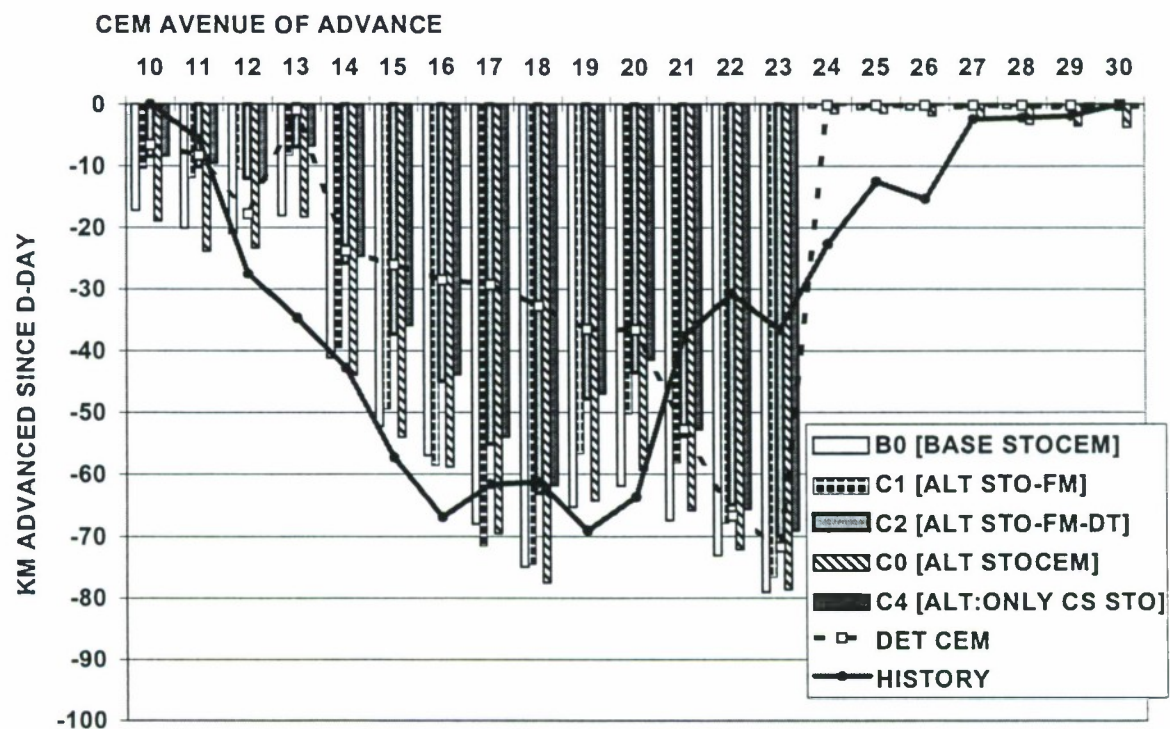


Figure E-3. FEBA Positions on D+12



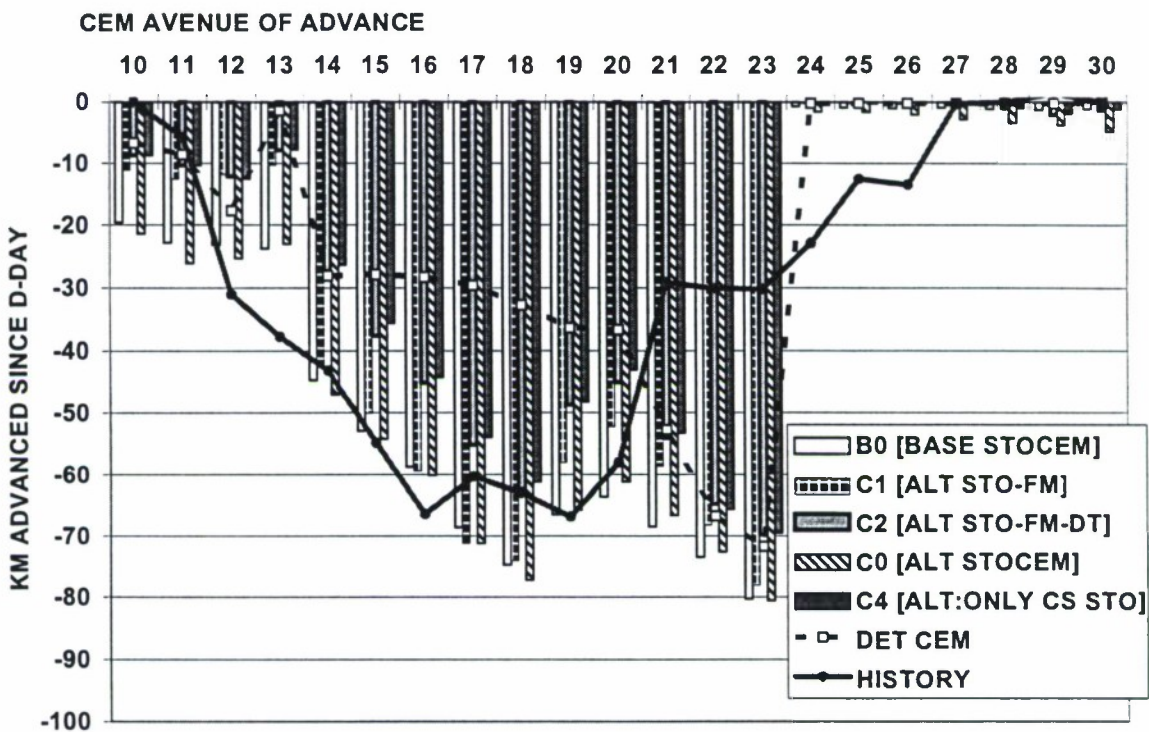


Figure E-4. FEBA Positions on D+16

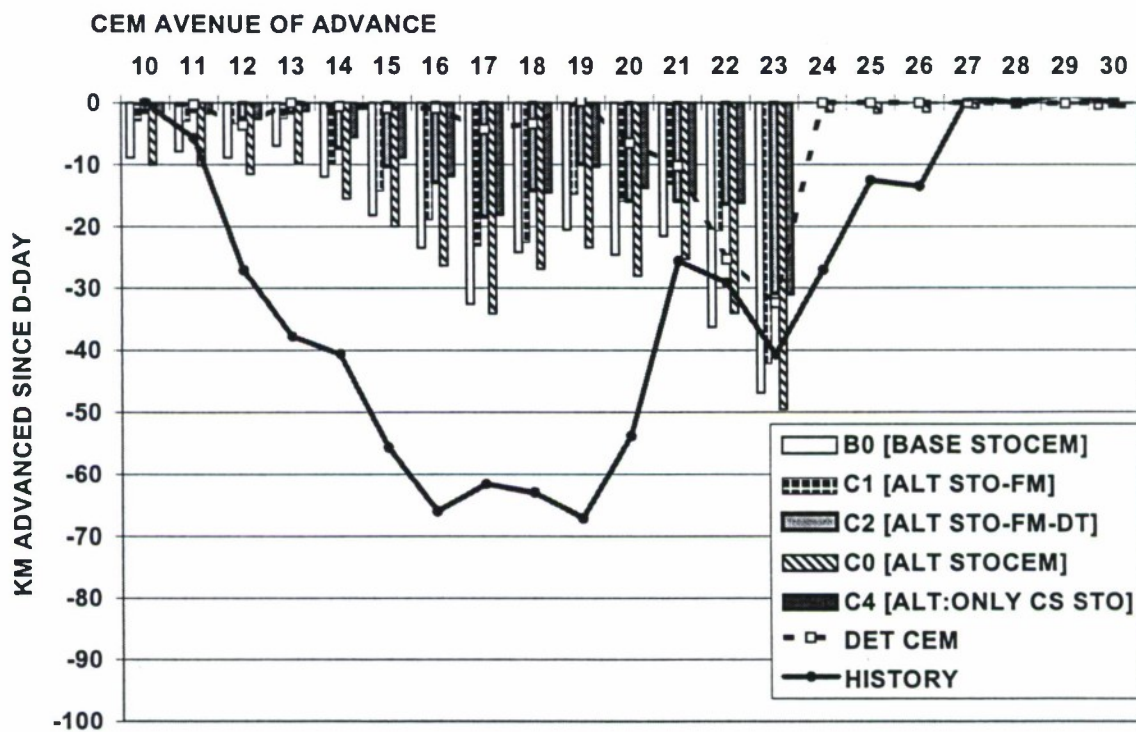


Figure E-5. FEBA Positions on D+20

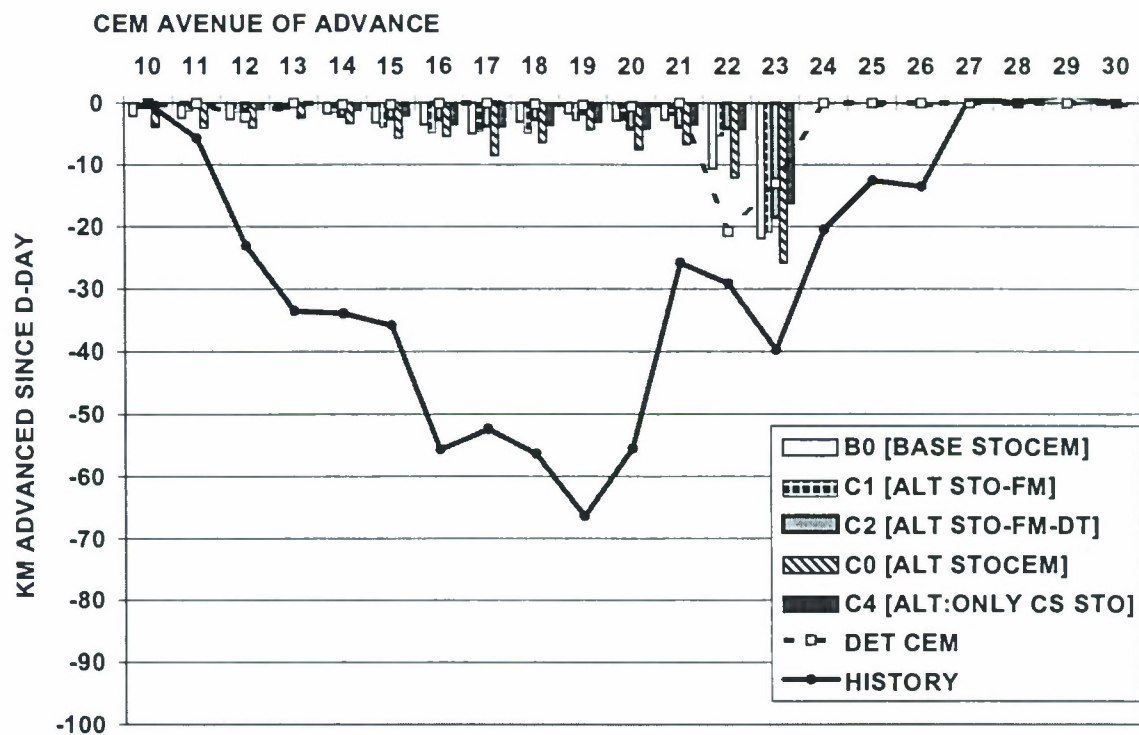


Figure E-6. FEBA Positions on D+24

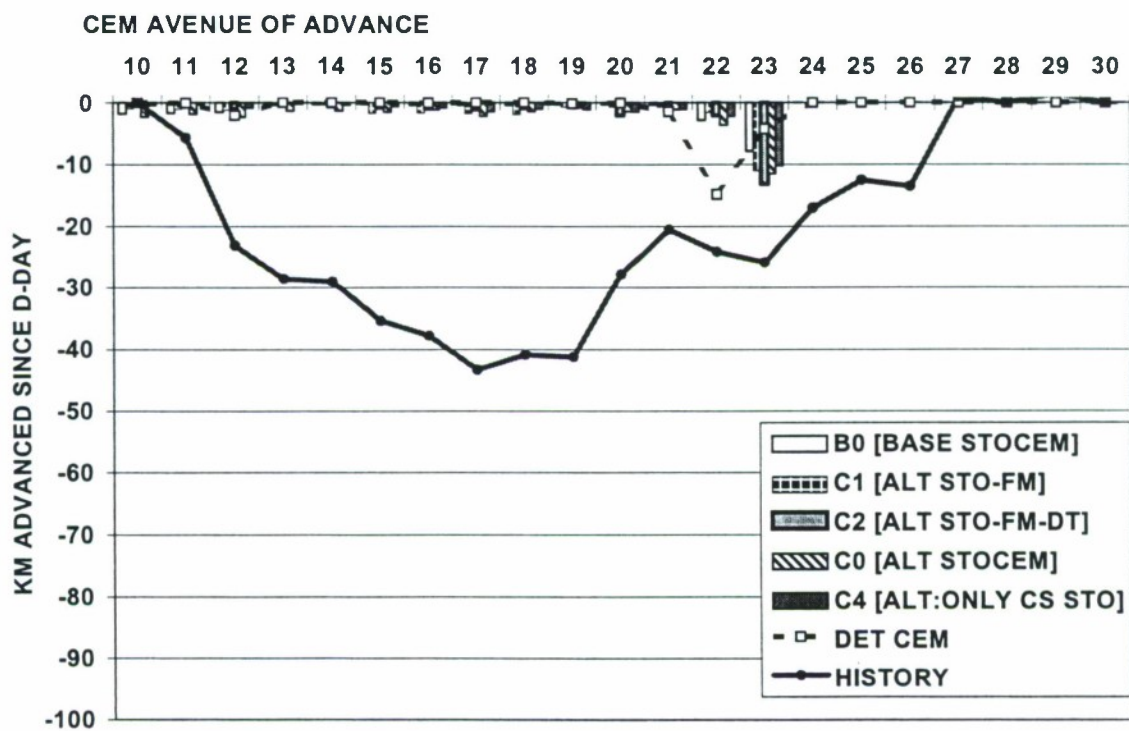


Figure E-7. FEBA Positions on D+28

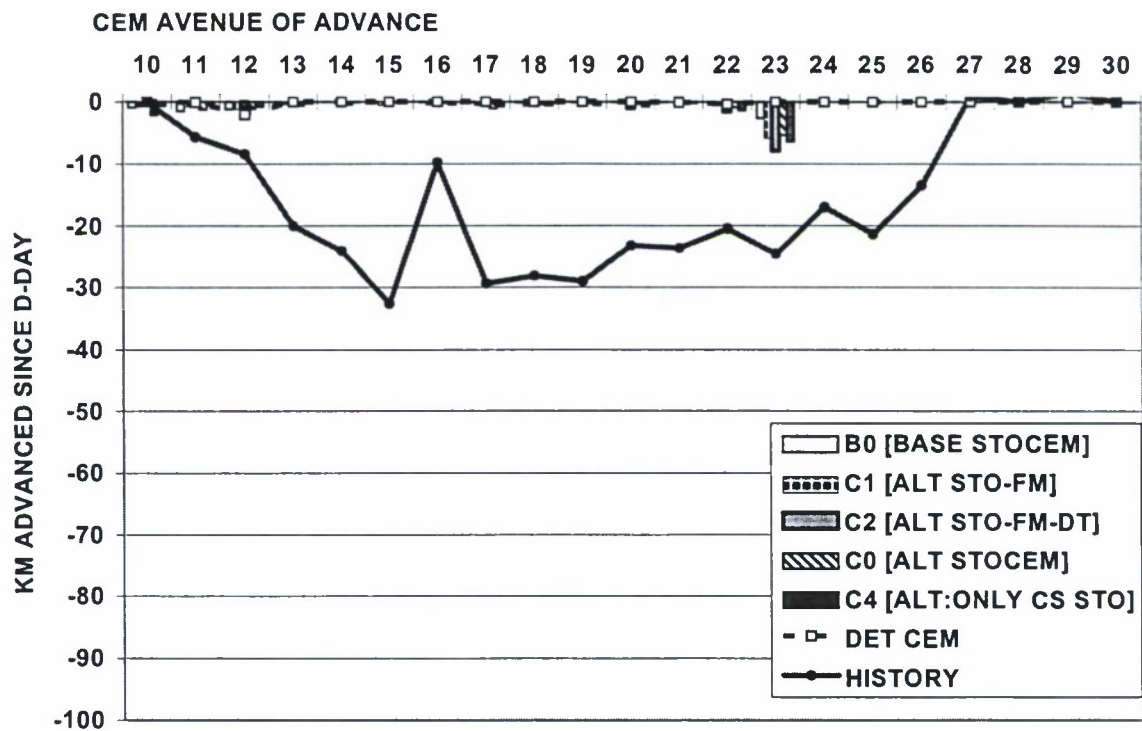


Figure E-8. FEBA Positions on D+32

## GLOSSARY

### 1. ABBREVIATIONS, ACRONYMS, AND SHORT TERMS

ACSDB	Ardennes Campaign Simulation Data Base
ADDCOP	Automated Data Display of CEM Outputs
ALT	alternative
ARCAS	Ardennes Campaign Simulation (study)
CAA	US Army Concepts Analysis Agency
CAS	casualties
CUM	cumulative
CEM	Concepts Evaluation Model (simulation)
CS	COSAGE sampling (stochastic process)
D-day	day 1 (of campaign scenario)
DET CEM	deterministic CEM (case)
DNBI	disease and nonbattle injuries
DT	decision threshold (stochastic process)
FEBA	forward edge of the battle area
FM	FEBA movement (stochastic process)
GE	German
KOSAVE	Kursk Operation and Simulation and Validation Exercise (study)
km	kilometer(s)
max	maximum
min	minimum
MOE	measure(s) of effectiveness
PCT	percentile
PER	personnel
RAA	research analysis activity



SEACA	Simulation Enhancements from the Ardennes Campaign (RAA)
SICS	STOCEM Investigation of Combat Sampling (RAA)
STO	STOCEM or stochastic
STOCEM	Stochastic CEM (simulation)
TC	theater cycle
Tk	tank
US	United States
UK	United Kingdom

## 2. MODELS, ROUTINES, AND SIMULATIONS

CEM IX	Concepts Evaluation Model IX - a two-sided, fully automated, deterministic model capable of aggregating conventional warfare results as a series of 4-day theater-level cycles
COSAGE	Combat Sample Generator - a two-sided, stochastic, high-resolution (division level) simulation model which simulates a day's combat activity to generate ammunition consumption and equipment and personnel loss data
STOCEM	Stochastic Concepts Evaluation Model - a stochastic version of CEM IX, a two-sided, fully automated model capable of aggregating conventional warfare results as a series of 4-day theater-level cycles